An Investigation of the Factors which Explain Variation of the Content of Sell-side Analysts' Reports

Nana Oiza Akubelem
BAcc, MRes

Submitted in Fulfilment of the Requirements for the
Degree of Doctor of Philosophy in Accountancy

Adam Smith Business School
College of Social Sciences
University of Glasgow
October 2015

Abstract

This thesis examines factors which may explain content variation in sell-side analysts' reports. There are two main objectives: (i) to ascertain whether the extent of accounting information contained in these reports varies with firm and analysts' characteristics; and (ii) to examine whether the tone and readability of the reports vary with analysts' incentives to produce optimistic research. Based on a sample of 288 reports on 144 S&P 500 firms, the first objective was addressed using a manual content analysis to examine accounting themes, while the second objective was addressed using automated content analysis based on context-specific and user-defined wordlists.

The empirical results indicate that the extent of use of accounting information in analysts' reports varies across firm characteristics but such variation only partly reflects its relevance for valuation as suggested by the value relevance literature. Moreover, analysts' incentives are influential as reports issued by analysts employed by investment banking firms or those in possession of the Chartered Financial Analysts qualification contain more references to forward-looking accounting information. Patterns of strategic reporting are also identified, as analysts employed by investment banking firms issue less readable reports compared to analysts employed by independent research firms following the same company. Further, readability is lower when the reports are less optimistic, indicating a tendency to obfuscate bad news through more complex reporting. Overall the findings are consistent with an impression management perspective as it reveals that content of analysts' reports may not be entirely objective but influenced by analysts' incentives to promote the companies covered.

The thesis contributes to extant literatures on the relevance of accounting information, content of analysts' reports, analysts' bias and impression management. Moreover, the findings have policy implications as they speak to the concern about the relevance of accounting information and highlight the need to consider the subjective influences and the role of analysts' incentives. Additionally, policy intervention on analysts' bias should extend beyond recommendations and earnings' forecast and consider the largely unregulated nature of the narrative content of the reports.

Table of Contents

List of Tables

List of Figures

Acknowledgement

First, I wish to first express my utmost gratitude to Almighty God for the grace to successfully complete this PhD thesis.

Second, I also wish to offer my sincere thanks to my supervisors, Professor Vivien Beattie and Dr Mark Aleksanyan for accepting me as a PhD student, believing in my abilities, their valuable support, guidance and patience through-out the course of my studies. I am very grateful for time spent in reading and commenting on several drafts of my thesis. However, any remaining errors and omissions are my sole responsibility.

Third, I wish to acknowledge financial support from the Adam Smith Business School scholarship fund, which enabled access to very relevant software programmes used in this research as well as funding for conference participation. I wish to also thank the participants at the Financial Reporting and Business Communication Conference in Bristol from 2012-2014 and Centre for Impression Management in Accounting Conference in June 2014, who offered constructive comments on various aspects of my research.

Fourth, I would like to express gratitude to academic members of staff of the Department of Accounting and Finance at the University of Glasgow, who at various times, provided some comments on my research or help identify resources useful for my research.

Finally, I am extremely indebted to my family and friends who have supported me throughout the course of my studies in various ways. Their love and constant support and prayers kept me going.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature

_____

Nana Oiza Akubelem

Chapter 1: Thesis introduction, background and motivation

1.1 Chapter Introduction

This chapter is aimed at introducing the thesis and providing an overview of the research by presenting the background and motivation for the study, the research objectives and related research questions and the structure of the thesis. The chapter commences with a discussion of the background and related information in section 1.2. The research problem and research objectives of the study are presented in section 1.3. Specific research questions and approach used to address these are outlined in section 1.4 and the overall conclusion and contribution of this research to extant literature is discussed in section 1.5. Section 1.6 discusses the motivation for engaging in this research. Finally, the outline of the thesis, including a brief summary of subsequent chapters is contained in section 1.7

1.2 Background

Financial analysts play a significant role in capital markets as information intermediaries. Their main tasks include gathering information from various sources about the companies covered, analysis and processing of the data and dissemination of the information in the form of earnings' forecasts, recommendations, price targets and written analysis (Michaely and Womack, 1999). Several studies attest to the importance of analysts' activities in enhancing the informational efficiency of stock prices as the speed with which stock prices reflect public information increases with analysts following (Hong *et al.*, 2000). Further analysts provide information incremental to that contained in corporate disclosures (Frankel *et al*. 2006) and they improve the information environment of the companies they follow in general, resulting in increased investor recognition, reduced information asymmetry, reduced cost of capital and improved trading and liquidity (Irvine 2003, Roulstone 2003, Bowen *et al*. 2008).

Corporate managers reckon that analysts are key influencers of their company's stock prices and are given considerable attention (Graham *et al.*, 2005). Moreover, their research is relied on by both institutional and retail investors (Boni and Womack, 2002; Campbell and Slack 2008). Beyond their role as information

intermediaries, analysts are perceived as sophisticated users of accounting information and thus a proxy for investors in general (Schipper, 1991).Given the relevance of analysts' activities and their sophisticated use of accounting and other information, they have been the focus of much academic research activity. Empirical research on the role and activities of analyst is vast. Bradshaw (2011) provides a recent review of the state of research on sell-side analysts and observes that despite several decades of academic research, there is very limited information on what analysts actually do, which remains a "black box". Bradshaw (2011) further argues that a means of understanding what analysts do is to examine the outputs of their research activity. Nevertheless, the bulk of existing academic research has focused exclusively on the quantified output of analysts' research activity such as earnings' forecast and stock recommendations and to a limited extent, target prices. "It is necessary to focus on other activities performed by analysts and attempt to better model their incentives than has typically been done" (Bradshaw 2011, p. 3). This call for additional research into other aspects of analysts work was earlier made in Schipper (1991).

1.3 Research problem, objective and questions

In responding to this call, this study seeks to examine analysts' written report in its entirety. Given the profound focus on analysts' earnings forecasts and stock recommendations in existing literature, there have been repeated calls to extend the knowledge of what analysts' do by examining other elements of analysts' research output beyond these quantitative measures, such as the written analysis contained in their reports (Schipper, 1991; Bradshaw 2011). However, studies on the written content of analysts' reports have been relatively few and the topics investigated very limited. In general, researchers investigate the types of information contained in the reports (Govindarajan, 1980; Previts *et al*. 1994; Rogers and Grant 1997; Fogarty and Rogers 2005; Flöstrand and Ström, 2006; García-Meca and Martínez, 2007; Abhayawansa and Guthrie, 2012), or the capital market reaction to the information contained in the reports (Asquith *et al*. 2005), or the linguistic features of the reports (Twedt and Rees, 2012; De Franco *et al*. 2015). However, it is argued that to advance our understanding of analysts' role and activities, it is imperative to investigate how and why information contained in the reports vary between analysts who cover the same company, or vary across

companies. Moreover, there is need to investigate the linguistic characteristics which reveal how analysts communicate information and why these features vary across analysts and across firms. This study seeks to extend knowledge of analysts through an examination of the narrative content of their reports in response to above mentioned calls and the gaps in extant literature on the content of analysts' reports. The specific research objectives of this research are summarised below:

- To examine the extent to which analysts' use a variety of accounting information in explaining and justifying value and stock recommendations in their reports. This extends existing literature which has mostly focused on reporting average values with little insight on how information in the reports actually varies. Moreover, a variety of accounting information is investigated which provides additional detail as to information types used by analysts.

- To investigate the factors that explain the variation in the use of a variety of accounting information in analysts' reports. This extends existing literature as it provides insight on whether usage of information is mostly driven by relevance for valuation as currently argued (Flöstrand and Ström, 2006) and provides opportunity to understand the impact of company, analysts and report characteristics on the observed variation in the use of accounting information. This is important because it improves our understanding of analysts' decision making process by uncovering the determinants of their choices of input (information types) to their decision making. What is more, knowledge of the factors which influence the extent of use of accounting information is important for policy makers as it enables them tailor responses to issues of irrelevance of accounting information accordingly.

- To examine how analysts communicate information in their reports (i.e. the linguistic features) and whether these linguistic features vary strategically in line with analysts' incentive to produce optimistic research. The linguistic features investigated in this study are the tone and readability. This extends the literature on the linguistic features, which has mostly been concerned about market impact, by providing evidence on whether tone and readability vary with analysts' incentives to produce optimistic research. This is important as it improves understanding of how analysts' communicate with investors in

their reports. Moreover, knowledge of whether analysts' incentive to produce optimistic research results in reporting bias of the written analysis in their reports is important for market regulators and policy makers as it provides relevant input into policy intervention aimed at addressing analysts' bias.

Two main research questions are thus addressed in this research: 1) what company-, analyst- and report-specific factors explain the extent to which analysts use accounting information in their reports? 2) Does the tone and readability of analysts' report vary with analysts' incentives to produce optimistic research?

1.4 Research approach

The first part of the research, which seeks to address the first research question involved manually coding and subsequent content analysis of 288 analysts' reports, using a detailed coding scheme. Content analysis was used to derive the dependent variable which measures the extent of use of accounting information in analysts' reports. The variable is measured as the frequency with which accounting information is referred to in analysts' reports.

For the purpose of this research, accounting information is defined as information contained in the financial statements, a sub-component of information contained in the financial statements or its uses in the form of accounting ratios, valuation models and financial forecasts. Additionally, references to accounting information are distinguished based on their time-orientation. The first research question is then addressed through a series of statistical tests and multivariate regression analysis.

For the second part of this research, which seeks to address the second research question, the first step involved deriving measures of tone and readability of analysts' reports. Tone was measured based on the frequency of occurrence of positive and negative words contained in a combined wordlist developed by Loughran and McDonald (2011) and Henry (2008) and readability was measured using a recently developed professional readability score, i.e. the *Bog index*.

Further, analysts' incentives to produce optimistic research are proxied by analysts' employer. A distinction is made between analysts employed by investment banking and brokerage firms (IB-analysts, hereafter) and those employed by independent research firms (IND-analysts, hereafter) which do not provide investment banking or brokerage services. Paired sample analysis is used to test whether tone and readability vary across these analysts, in accordance with their incentive to produce optimistic research.

1.5 Overall conclusions and contribution

With regard to the first research objective, the findings in this study provide evidence that the extent of use of accounting information varies with some company characteristics. The direction of the association between company characteristics and extent of use of accounting information in the reports partly departs from those expected if analysts' were principally concerned with valuation (based on the value relevance literature). Consequently, the study provides input towards future theorising of the rationale behind analysts' use and disclosure of information in their reports. Moreover, the study provides sub-level categories of accounting information and distinguishes between references to historical and forward-looking information. The results reveal that the extent of use of different categories of accounting information and across different time-orientations differ significantly and provide additional information beyond the broad level categorisation used in previous studies. Such lower level distinctions reveal that analysts use a broader range of information items than those deemed relevant under the value relevance models.

With regard to analyst-firm characteristics, it was found that analysts employed at investment banks use more accounting information and rely more on forward-looking information than do analysts who work at independent research houses. This suggests that the former group of analysts either have superior forecasting ability, more access to superior information useful for forecasting, are more sceptical of historical accounting information or intentionally focus on discussing forward-looking unaudited estimates in order to provide more optimistic outlook for the companies covered. Furthermore, the difference in the use of accounting information also varies across analysts with or without the CFA designation. These distinctions have been largely ignored in prior literature, which has focused mostly

on distinguishing the output of different types of analysts, rather than understanding how their differences influence analysts' accounting information processing.

With regard to the second research objective, the findings suggest that the narrative content of analysts' reports is important when examining analysts' bias. The main difference between IB and IND-analysts linguistic features are in the use of complex reports (i.e. reports with low readability). As documented in prior studies (such as Li, 2008), increased complexity may be strategically used to obfuscate bad news. IB-analysts' strategy of reporting bad news principally involves the use of more complex reports relative to IND-analysts and the results suggest that the variation in the readability of IB-analysts' reports is systematically associated with the level of optimism of the reports as measured by the tone. IB-analysts' reports are more readable when tone is higher and less readable when the tone is lower. The pattern is not observed for IND- analysts which are arguably (or at least perceived to be) more objective.

Although the general pattern is that the tone of IB-reports is not significantly different from those of IND-reports following the same companies, when the sample is limited to companies for which IB-analysts have issued relatively positive recommendations, both the tone and readability of the reports are significantly higher than the IND-reports on the same companies. The same pattern is not observed for IND-analysts. This provides a useful extension of prior research on analysts' bias and suggests that analysts' incentives to produce optimistic research is also reflected in the narratives which accompany their forecasts and recommendations.

1.6 Research motivation

The research questions and particularly the focus on analysts' reports is motivated by a number of factors. First, the call to extend knowledge of what analysts do by examining other elements of analysts' research output (Schipper, 1991; Bradshaw 2011). Analysts' reports provide a rich source of information about what analysts do to necessitate academic enquiry. Analysts' reports represent a medium of communication of analysts' opinion to existing and potential investors. They contain factual information about the companies, analysts' own judgements

(Abdolmohammadi *et al*. 2006), valuation models used to derive the target prices, (Demirakos *et al*. 2004) and information about peer companies. Given the use of the reports as a means of reporting/communicating with investors, analysis of the linguistic attributes of the reports provide a rich source of information into "how" analysts communicate information. Moreover, the content of analysts' reports is mostly discretionary as there are no mandatory requirements to disclose specific types of information. Thus, an analysis of the variation of the content of the reports provides a relevant step toward understanding what analysts' do and addressing the calls for more research in this area.

Second, a number of empirical studies provide evidence that the content of analysts' reports contain additional price-sensitive information beyond quantitative measures such as earnings' forecasts and recommendations (Asquith *et al.*, 2005, Twedt and Rees, 2012 and De Franco *et al*. 2015). Consequently, the reports represent a significant piece of research output that should not be ignored if we are to extend our knowledge of the work of analysts.

Third, in a review of the literature on the corporate information environment, Beyer et al. (2010, p.335 emphasis added) conclude that "understanding what role ….the incentives analysts face, play in determining the decision to follow a firm and not simply when to issue a report but **what to include in that report** is an important avenue for future research to deepen our understanding of the development and effects of the corporate information environment". This research addresses this need to develop existing literature along this line as it investigates among others, the association between analysts' incentives and their use of accounting information as well as the linguistic feature of their reports.

Fourth, research in social psychology and studies of impression management within the context of corporate reports suggest that choice of words reveals important information about the authors. Several linguistic features of corporate reports have been examined and evidence from these studies suggests that linguistic features vary across company characteristics in a manner that serves managerial interest (Clatworthy and Jones, 2003, 2006; Li, 2008; Merkl-Davies *et al*. 2011). An extension of a study of word choice and linguistic attributes to the narrative content of analysts' reports provides a useful means of extending previous research on analysts' bias. To date, the research on analysts' bias has

focused principally on quantitative measures such as earnings' forecasts and recommendations rather than the accompanying narratives. Given market regulator's interest in addressing analysts' bias, an understanding of how analysts may bias the narrative content of their reports provides relevant input for the decision making of policy makers and capital market regulators interested in addressing analysts' bias. A study of the narrative content of analysts' reports and variation in the linguistic features is therefore topical and relevant to a broad range of parties. Analysts may project optimism in their written reports in a number of ways beyond the strategic selection of information. These may include praising managerial actions and failure to criticise managerial decisions (Fogarty and Rogers 2005, Westphal and Clement 2008), optimistic tone of reports, obfuscation of bad news through complex reporting, etc. Hence, this study sheds more light by investigating how the narrative content of the reports vary and what motivates the variation.

1.7 Structure of the thesis

The thesis comprises two parts and contains 11 chapters in total. Chapter 1 has introduced the background, motivations, research objectives and research questions. Given the focus of this research on sell-side analysts, Chapter 2 provides an overview of the role and activities of this group of professionals. The chapter includes a discussion of the analysts' activities and their influence in the capital markets. Additionally, a description of the institutional setting in which they operate, including employers, compensation structure, conflicts of interests, regulatory setting which govern their activities are also discussed. The distinction between IB and IND-analysts is central to this thesis. Hence, the differences between these two groups are extensively reviewed. This is followed by a description of analysts' reports, which is the subject material of analysis in this research. This serves to provide relevant background information that will enhance understanding of the whole research process.

Research objective 1 is addressed in Chapters 3-6, while research objective 2 is addressed in Chapters 7-10. Chapter 3 reviews existing literature on the use of accounting information by analysts, including survey based, experimental and content-analytic studies. The gaps in existing literature are established and the research objective is restated.

Chapter 4 presents and discusses the hypotheses aimed at testing whether company-specific, analyst-specific and report-specific variables explain the variation in the use of accounting information in analysts' reports. Value relevance literature is relied on as the basis for hypotheses formulation. Furthermore, proxies for the independent variables and the statistical model used to test the stated hypotheses are also presented.

Chapter 5 provides a detailed description of the methodological approach used to derive the dependent variable which measures the extent of use of accounting information. Manual content analysis of analysts' reports was conducted. Choices made with regard to the content analysis design are discussed and justified. These include decisions on the recording and measuring units, categories of accounting information, development of a coding instrument, data collection and tests of reliability. The chapter also discusses the sample selection process, details of the data and sample composition.

Empirical results are presented and discussed in Chapter 6. The chapter begins by presenting descriptive statistics of the dependent and independent variables used in this study, followed by tests of pre-stated hypotheses and related discussion. Robustness tests conducted are also discussed.

Chapter 7 reviews existing literature on the tone and readability of analysts' reports. The gaps in existing literature are established and the research objective is restated. This is followed by a review of the literature on analysts' conflicts of interest and empirical evidence of the difference between IB and IND-analysts' forecasts and recommendations is presented. Social psychology literature on the relevance of word choice is reviewed as well as its application to the study of corporate narratives. Finally, the theoretical framework of ingratiation is used to explain the possible influence of analysts' employment structure on the linguistic features of their reports.

On the basis of the studies reviewed in Chapter 7, Chapter 8 presents and discusses the hypotheses aimed at testing whether analysts' employment structure (which is used to proxy for the incentives to be optimistically biased) explains variation in the linguistic features of analysts' report. The research design used to test the hypotheses is also stated and justified.

Chapter 9 provides detailed description of the methodological approach used to derive measures of tone and readability for this study. For the tone variable, a dictionary based approach based on the frequency of occurrence of positive and negative words is used. The dictionary used to grade the tone of the reports is based on a combined list of Loughran and McDonald (2011) and Henry (2008). The choices made are justified and tests of validity of the wordlists conducted. For the readability measure, the *Bog index* is used. This is a recently developed measure of readability which captures several aspects of syntactic complexity that the traditional measures such as the Fog and the Flesch indices ignore.

Empirical results are presented and discussed in Chapter 10. The chapter begins by presenting descriptive statistics and univariate analysis, followed by tests of pre-stated hypotheses and related discussion. Robustness tests conducted are also discussed.

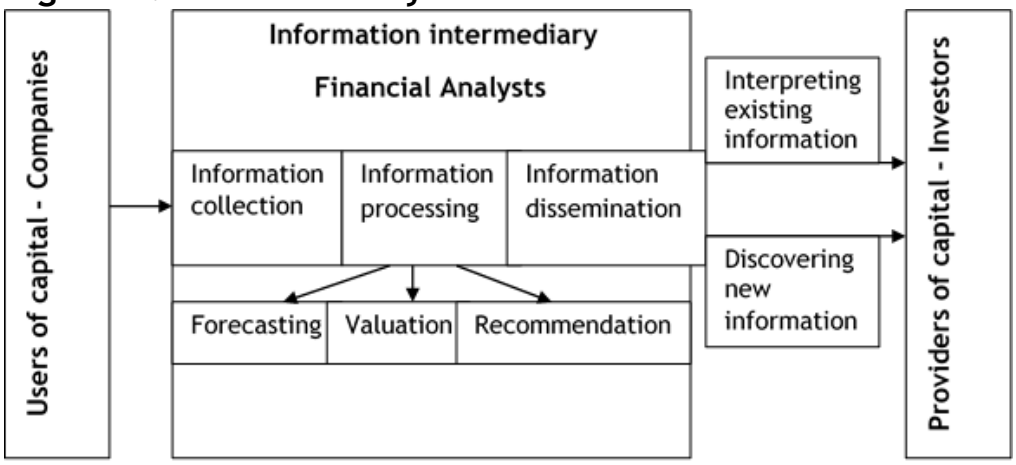
Chapter 11 concludes the thesis, providing a summary of the research, main contribution to existing literature, the implications of the findings, limitations of the research and avenues for future research.

Chapter 2: Background Information on sell-side analysts

2.1 Sell-side analyst's role

Ljungqvist et al. (2007, p. 421) argues that "A sell-side analyst's primary role is to act as an informational intermediary, channelling information (in the form of investment recommendations, earnings forecasts, and detailed reports) from companies to investors". Theoretically, within the capital market, analysts function as information intermediaries between (companies) the users of capital and (existing and potential investors) the providers of capital (Jensen and Meckling 1976, Fisch and Sale 2002). They are key players in the market for information[1] in general (Holland and Johanson, 2003) and the information supply chain[2] in particular (Campbell and Slack, 2008). On the one hand, they represent the demand-side as they gather information from companies and other sources and on the other hand, they represent the supply side as they disseminate information about companies to investors. Liu et al. (2007 p. 630) states that "In general, analysts add value by both (1) aggregating publicly available information and (2) generating new information." Figure 1 provides a depiction of this function.

Figure 1: Sell-side analysts' role



[1] Holland and Johanson (2003) defines the 'market for information' as the institutional means to connect corporate information supply activities to security market information demand activites.

[2] Campbell and Slack (2008, p. 3) describe the 'information supply chain' as the process of production of accounting information (by preparer companies), its interpretation (by sell-side analysts) and its consumption (by buy-side).

As depicted in Figure 1, analysts intermediate between companies and investors, playing two distinct yet related roles, i.e. that of information discovery and information interpretation (Chen et al. 2010[1], Livnat and Zhang 2012). The information discovery role consists of analysts' ability to add value to the investment process through extensive search for relevant private information from different sources[2], which is then incorporated into their forecasts and recommendations to investors. The interpretation role depends on analysts' ability to employ their superior knowledge and expertise in the analysis and interpretation of public available information. Their activities begin by collating information from various sources about the company being followed. Employing their financial expertise and knowledge, such information is analysed, processed and used for forecasts of future earnings and other accounting metrics and value drivers. These serve as input into valuation models for estimation of a stock's intrinsic value. The resulting output is an investment recommendation, accompanied by written reports (Michaely and Womack, 1999) which are disseminated to investors.

Analysts are highly influential in the capital markets and their activities offer several benefits both for companies and investors (Bradshaw 2011). As managers have incentive to misrepresent information (the famous lemons problem – Akerlorf, 1970), analysts' intermediation role enables external monitoring of managerial behaviour, thus enhancing the credibility of corporate disclosures (Yu 2008, Irani and Oesch 2013). Moreover, given their expertise, industry knowledge and regular interactions with corporate managers, analysts are able to scrutinise managerial decision making, track company performance over time and compare them with peers while also probing any unusual observations. In the theory of the firm, Jensen and Meckling (1976) argue that analysts reduce the agency cost associated with separation of ownership and control and are therefore socially productive. Empirically, Doukas (2000) found a negative association between proxies of agency cost and number of analysts' following. Similarly, Yu (2008) found a negative association between number of analysts following a company and earnings management practices. Chen *et al.* (2012) extends Yu (2008) by observing cash holding, CEO compensation and acquisition decisions and provides additional support for the claim that analysts' coverage reduces managerial opportunism. Also Weirsema and Zhang (2011) provide evidence that analysts' negative

recommendations on a company are associated with increased likelihood of CEO dismissal.

In general, the importance of analyst's role as documented by their positive impact on capital market operations and the reliance on their research by investors (Campbell and Slack, 2008) makes them an interesting object of academic research and motivates this research.

2.2 Activities of financial analysts

Although analysts' principal role in capital markets is that of intermediation, their activities are multifaceted (Fernandez 2001).
The Securities and Exchange Commission's regulation 'Analysts' Certification' (SEC, 2003) defines a research analyst as "any natural person who is primarily responsible for the preparation of the content of a research report" with research report defined as "a written communication ... that includes an analysis of a security or an issuer and provides information reasonably sufficient upon which to base an investment decision". This indicates that analysts' main activity is to conduct company research (Abhayawansa, 2010) which can be distinguished into three phases as depicted in Figure 1: i) information gathering; ii) information processing and iii) information dissemination. These are discussed below:

Information gathering

To form an opinion about the value and prospects of the companies covered, analysts need information. Thus, the first stage of their research process involves collating of relevant information from various sources, principally from corporate managers (Fogarty and Rogers, 2005). Information gathering from companies involve reliance on publicly disclosed data such as annual reports and other SEC filings (Abdolmohammadi *et al*. 2006) and private information gleaned from the frequent interactions with corporate managers. Such interactions are crucial to the work of analysts, they provide useful input into analysts' earnings forecasts and recommendations (Brown *et al*. 2015; Soltes 2014) and are perceived as an important channel of communication by corporate managers (Marston 2004). They may involve face-to-face meetings either privately or in groups such as earnings

press releases and conference calls, analysts' days and brokerage organised conferences[3] (Kirk and Markov 2012; Green et al. 2014). Interactions may also take the form of phone calls and email messages (Barker 2000, Marston 2004). Brown *et al*. (2015) provide recent survey evidence which reveal that analysts' interactions with corporate managers still represent an important source of information despite the enactment of regulation fair disclosure (Reg FD hereafter) which inhibit selective disclosure of material non-public information to stakeholders such as sell-side analysts. Interviewees in that study reveal that such interactions now provide avenue for analysts to seek understanding of public disclosures by asking follow-up questions, requesting explanation of performance and judging managerial optimism through their vocal cues and body language. Thus, information gathering by analysts is not limited to hard quantifiable data. Beyond corporate sources, analysts also rely on external stakeholders for information including suppliers, customers, trade groups etc.

Information processing

The second stage of analysts' research activity involves processing the information gathered, in order to form an opinion about the company's values and prospects (Fernandez, 2001). As indicated in Figure 1, the processing and analysis of information includes forecasting and valuation, the outcome of which is a recommendation to investors. The process by which information is used to forecast earnings and estimate value and how these values are used to arrive at investment recommendations is unobservable or a "black box" (Bradshaw, 2011). However, a number of prior empirical investigations have shed some light into how analysts process information. First, analysts use a wide variety of information but their relevance is "temporal and contextual" (Abhayawansa et al. 2015, p.4). Both Bouwman (1995) and Barker (2000) document that earnings related information only dominate the early stages of investment analysis. Moreover, usage of

[3] Earnings press releases are quarterly announcements of companies' financial results. Conference calls are non face-to-face interactions between company executives and stakeholders such as sell-side analysts and institutional investors, which occur typically following quarterly earnings' announcements. Analysts' days are face-to-face disclosure events organised and hosted by companies and are aimed at providing in-depth information to equity analysts and institutional investors. Brokerage organised conferences are invitation-only events organised and hosted by brokerage firms and are aimed at "connecting select institutional clients with company management" (Green, et al. 2014, p.143).

information varies across industry (Matsumoto et al. 1995; Coleman and Eccles, 1997; Abdolmohammadi *et al*. 2006), recommendation types (Breton and Taffler, 2001), leverage and stock return volatility (Orens and Lybaert, 2010) and direction of financial performance (Coram *et al*. 2011).

Second, analysts do not use different information types in isolation. For instance, usage of accounting information is complemented by non-accounting information such as information on the quality of management, corporate strategy and growth (Barker and Imam, 2008). Third, forecast of earnings, estimate of value or investment recommendations are not deterministic. Analysts derive most of these values based on heuristics (Bradshaw 2002, 2004) and rely significantly on qualitative information through-out the process, such as intellectual capital information (Abhayawansa *et al*. 2015). This implies that analysts information processing is proprietary and though analysts may be presented with the same information about a company, they may process it differently and arrive at different conclusions. This demonstrates the relevance of understanding analysts' characteristics and how it may impact on what analysts' do.

Information dissemination

As depicted in Figure 1, analysts research activity results in discovery of private information and interpretation of existing public information. As intermediaries, analysts acquire and process information about the companies they cover and disseminate such information to investors. Such information is summarised and disseminated to analysts in the form of earnings' forecasts, price targets, recommendations and written analysis contained in their research reports. Extant empirical studies demonstrate that these research outputs have information content and are value relevant (Asquith *et al*. 2005; Frankel *et al*. 2006). Analysts also disseminate information informally such as through private phone calls with the buy-side or through the business media. Given that analysts' reports are the object of this study, the next section provides a description of the reports.

2.3 Analysts' reports

It is well recognized that one of the outcomes of financial analysis by equity analysts is the production of a research report in which analysts discuss their investment recommendations, provide forecasts of the firm's fundamentals, price target and a justification of this. While the actual process of analysis is hidden from investors, the research reports produced can be a "lens" through which the "black box" of analysts can be understood.

Analysts' reports are a formal channel through which sell-side analysts disseminate the output of their research activities (Campbell and Slack, 2008). Analysts' reports are principally used by institutional investors, who rely on these for asset allocation and portfolio management (Umiastowski, 2014). A number of empirical studies attest to the important value of sell-side analysts' reports to the buy-side (See for instance, Vergossen, 1993; Moizer and Arnold, 1984 and Holland 2006). It is generally recognised that although buy-side analysts rely on a number of sources of information, "they tend to rely quite heavily on sell-side analysts' reports" (Campbell and Slack, 2008, p. 6). Moreover, there is evidence of significant market reaction to the thematic (Asquith et al. 2005) and linguistic content (Twedt and Rees, 2012; De Franco *et al*. 2015, Huang *et al*. 2014) of the reports over and beyond the recommendations and earnings' forecasts.

Unlike annual reports which contain specific sections, the structure of analysts' reports vary considerably across and often within analyst-firms and across companies. Some analyst-firms (e.g. Morningstar) use a standard structure with headings, which are present in all the reports issued by analysts within the same analyst-firm, while other analyst-firms adopt a more flexible approach with reports structure that vary significantly across companies covered and across analysts. Usually the reports contain a summary page, which is followed by the main body and a regulatory disclosure section. The summary page is the first page and contains brief information about the company covered and a summary of the key points discussed in the main body of the report. Information disclosed in this area include company ticker, market data and relevant financial and valuation metrics such as current stock price, trading volume, dividend and cash flow yield and number of shares outstanding. These are generally presented in tabular format and in some cases contain a graphical display of share price or returns

performance over time. The stock recommendation, price target and EPS forecast are also disclosed in this section of the report. Information about the reporting analyst such as their names (accompanied by a designation which indicates their qualifications such as CFA, PhD, ACA) and their contact details, usually an email and telephone number are also displayed. The first page also contains highlights from the main report or an executive summary.

The structure of the main body of the reports vary considerably across analysts-firms and within analyst-firms. In general, most common sections that can be found in the reports are:

Events/Review section where the event which triggers the issuance of the reports are discussed. For this thesis, the results reports are used, which are reports issued just after the earnings' release. Hence, the events section contain description of the recently released annual results and key takeaway from this, including a comparison of actual results with analysts' consensus or managerial estimates.

Investment thesis section which presents analysis opinion about the performance and prospect of the companies covered. Usually contains the recommendation and a justification of this.

Risk section in which analysts highlight factors which may affect the performance of the company or impede the achievement of the price target. Some reports contain a scenario analysis where different possibilities are envisaged and impact on financial and valuation metrics discussed.

Company profile section in which analysts provide a brief description of the business of the company.

Valuation section in which the price target is disclosed, accompanied by the valuation method used to derive the figures and may also include disclosure of the estimates of accounting information used, e.g. earnings forecasts.

These sections are most commonly found but are not exhaustive as few reports also include additional sections such as a *Financial Strength* section or a

Management section which contain information about the company's debt profile, liquidity/solvency and information about senior management and corporate governance in general.

The final section of most of the reports features an appendix containing regulatory disclosures. Generally this consist of information about existing or potential trading or investment banking relationship with the company covered in the reports, analysts' source of compensation, analysts' certification that the views are accurate reflection of their personal opinions on the company, distributions of stock recommendations across buy, hold and sell and various disclaimers. These disclosures are generally lengthy, particularly for investment banking and brokerage firms and in general seeks to address the reporting requirement of various national and international regulations.

Beyond the structure, the content of the reports also vary considerably and information is presented in narrative, tabular and graphical forms and contains analysts' own assessment of the company, accompanied by factual information. Evidence from prior research suggest that the reports contain a wide range of information including accounting information (Fogarty and Rogers 2005, Abdolmohammadi et al. 2006), non-financial information such as intellectual capital information (Abhayawansa and Guthrie, 2012; Flostrand and Strom 2006, Garcia-Meca and Martinez 2007). Analysis often include discussion of the value drivers, justification of recommendations and detailed analysis such as SWOT analysis. Often analysis is reported by segments and also include peer comparison. Overall, the structure and the content of the reports vary significantly as analysts have discretion on the type of information presented in the reports and how this is presented.

Unlike the annual reports, analyst reports are issued several times a year following a variety of company events with the most common being the announcement of the quarterly and annual results. Timing and number of issuance vary significantly across analyst-firms. In general, there is very little research on the process of writing, editing and publication of the reports, influence of brokerage firms' policy or the extent of discretion used by analysts in the choice of content. However, Campbell and Slack (2008) provide some insight based on interviews with several analysts as they observed that the editing process often involves

submission of the reports to investors' relations department of the companies being covered for check of factual accuracy. Nevertheless, analysts have discretion on what facts to disclose in their reports as well as how to present these as documented in Beunza and Garud (2007). These authors also argued that analysts are frame-makers who provide a lens through which investors view the companies they report on. Given the variation in both content and format of their reports, there is need for research to uncover why analysts' reports vary in terms of information content as well as linguistic style. Flostrand and Strom (2006) argued that information is included in the reports if it is relevant for valuation while Umiastowski (2014) suggests that the reports mostly serve as marketing pieces for analysts who are trying to sell a story. Consequently, opinions on the company and therefore reporting choices and styles may vary. This thesis further provides empirical support for these arguments by identifying the factors which explains the variation of the content of analysts' reports.

2.4 Financial analysts: institutional and regulatory setting

2.4.1 Types of analysts

Equity analysis is carried out by research analysts who may be identified as buy-side or sell-side analysts. In this study, sell-side analysts are further distinguished into those employed by investment banks and/or brokerage houses (IB-analysts) and those employed by independent research firms (IND-analysts).

Buy-side analysts: are analysts employed by institutional investors such as insurance companies, pension, mutual and hedge funds. They provide equity research primarily for the benefit of their employers. Hence, their research is proprietary and not disseminated to the public. Due to the large portfolio of stocks they cover, they rely on sell-side analysts for input to their investment decisions.

IB-analysts: these are analysts employed by brokerage houses and investment banks and issue equity research reports that are disseminated to existing and potential investors. These analysts usually follow six to seven companies in a sector and provide support for the sales force in their banks (Davidson, 2008). They spend considerable effort researching on specific companies and stocks, unlike their buy-side colleagues, who focus more on the portfolio of stocks than individual stocks.

(iii) Independent analyst (IND): these are analysts employed by research firms which do not engage in investment banking or brokerage business. Like IB-analysts, their research output is disseminated to existing and potential investors.

This thesis examines only reports issued by IB and IND-analysts. The distinction between these two types of analysts is central to both research questions addressed in this thesis. In the first research question, test of whether reports issued by IB and IND-analysts on the same company differ in their use of accounting information and in the second research question, the differences in linguistic features across reports issued by IB and IND-analysts are explored. To provide relevant context for understanding the results and explanation of empirical tests

presented in this thesis, the following differences between IB and IND-analysts are highlighted:

Based on employer

IB-analysts are employed by investment banks and brokerage houses, which offer other investment services beyond equity research such as underwriting new stock issues, securities trading, investment management, corporate advisory etc. These firms tend to be very large with operations in several countries. On the other hand, independent analysts are employed by independent research firms which do not have investment banking or brokerage businesses and are relatively smaller in size (Jacob *et al*. 2008). Clarke et al. (2011) show that on average independent research firms existing before the Global Settlement in 2003 employ on average of 23 analysts while the new firms set-up following the settlement employ only 8 compared with investment banking firms which have large equity departments.

Sources of Research Funding

For investment banking and brokerage firms which employ IB-analysts, equity research departments are cost centres and are subsidised with revenue generated from other businesses (Sirri, 2004). Prior to the Global Settlement Act, research used to be funded by revenue generated from investment banking business and brokerage services. IB-analysts employed by these firms were thus engaged not only in equity research but proactively involved in marketing campaigns to procure investment banking deals. IB-analysts compensation was in part linked to revenues from these businesses. However, following analysts' regulations (discussed subsequently) which prohibited the funding of research with investment banking revenues, equity research by IB-analysts is now principally funded with revenue from trading commissions (Groysberg, 2013). Research reports are not directly sold to investor but cost of research is paid for through soft dollars (i.e. mark-up on trading commissions). On the other hand, as independent research firms do not engage in other businesses, research is directly funded with fees obtained from investors. In these firms, research is sold on a subscription basis and this constitutes their main source of income (IOSCO, 2003). Consequently, IB-analysts are better compensated by their investment banking employers compared to IND-

analysts. Boni and Womack (2002) report a managing director of research at an investment bank stating that "..Independent research firms will never be able to retain and compete for the best analysts: they will never be able to offer compensation packages that investment banking firms can" (p.27)

Quality of Research

Given the differences between IB and IND-analysts, several studies have investigated whether the quality of research produced by IB-analysts differ from those of IND-analysts. On the one hand IB-analysts are faced with several conflicts of interest as a result of their employer's Investment banking and brokerage businesses. Consequently, the objectivity of their research is undermined. On the other hand, given their size and resources, investment banks and brokerage houses are able to recruit highly successful and reputable analysts (Clarke *et al*. 2011) and provide access to more information channels than is available for IND-analysts (Jacob et al. 2008). Thus IB-analysts have an incentive to protect their reputation and information advantage which enhances the quality of research provided. Empirical studies have sought to examine the difference in information content of their recommendations, the distribution of their recommendations and whether there is a difference in the level of optimism and forecast bias between IB and IND-analysts. Overall, the evidence till date is mixed and is discussed extensively in Chapter 7 of this thesis.

Sources of information

Unlike their IND counterpart, IB-analysts have access to more information sources. Given the size of their employers, they have access to large resources to enable more proprietary research including conducting surveys (Jacob *et al*. 2008). They also tend to have greater access to management due to organisation of meetings such as investors' conferences and field trips, which are funded by their employers.

Overall, these distinctions reveal that although both analysts conduct equity analysis and their research is disseminated to investors, they differ in several ways that could significantly impact the output of their research activities. First, given the other services offered by IB-analysts' employers and the indirect involvement

of IB-analysts in these activities, they face conflicts of interest that interfere with the objectivity of their research (IOSCO, 2003). Second, the larger size and greater resources available for investment banking and brokerage firms suggest they may be able to attract more reputable analysts and provide access to wider information channels than Independent research firms. These differences are likely to result in differences in research output. The academic evidence hitherto is not only inconclusive but has focused mostly on recommendations and earnings forecasts and largely ignored other content of the reports which accompany these measures such as the narrative analysis. Hence, this research addresses this gap in the literature by examining whether differences in the content of analysts' reports can be attributed to different characteristics of analysts as discussed above as well as other company characteristics.

2.4.2 Analysts' conflict of interest

There are several allegations that the work of analysts is not entirely objective but plagued with conflicts of interest which predispose them towards providing optimistic research output for the companies they cover. In this section, the two most cited sources of alleged conflicts are briefly discussed while a detailed review of the empirical literature on analysts' bias is presented in Chapter 7 of the thesis. Sources of analysts' conflict of interest include:

Analyst's access to managers for information

Interactions with corporate managers represent an important source of information for analysts (Chen and Matsumoto 2006, Soltes 2014) and most of the information contained in analysts' reports are sourced from the firm (Fogarty and Rogers, 2005). Access to information from managers, through invitation to participate in conference calls, availability for private meetings and response to phone calls, is therefore crucial in obtaining superior insight about the future prospects of the firms followed (Mayew 2008; Brown *et al.*, 2015). Such dependency on managers for access to corporate information compromises analysts' objectivity as they are incentivised to issue reports which serve corporate managerial interest in order to maintain close relationships. This is achieved through issuance of optimistic stock recommendations, optimistic or

beatable earnings' forecasts and avoidance of criticism of managerial actions (Fogarty and Rogers, 2005). Analysts' who fail in this respect could face reprisals in the form of denial of access to private meetings, phone calls or discrimination in Q&A sessions (Campbell and Slack 2008, Mayew 2008).

The benefits to the analysts from private interaction with managers go beyond the acquisition of information. In a study of private interactions between managers and analysts, Soltes (2014) found that analysts seek access to corporate managers in order to provide "corporate access" to their clients. Analysts' facilitate meetings between corporate managers and their institutional clients by arranging for this meeting and covering the costs. Corporate access is valued by the buy-side (Valentine, 2010) and facilitating access to corporate managers is a key attribute accessed as part of the institutional investor rankings. Overall, personal benefits for the analysts of maintaining good rapport with corporate managers consisted of informational advantage that could lead to greater quality of their research output and also enhance their reputation with institutional investors both as a result of the quality of their research and corporate access.

Analysts' involvement in attracting investment banking business

Analysts are employed by firms which engage in various business activities beyond equity research, such as investment advisory, underwriting of securities and stock trading. Generally, analysts' research is subsidised/funded by these activities (Fisch and Sale, 2002), which creates the incentive for analysts to issue reports which help generate revenues for these businesses. For analysts-firms with underwriting and corporate advisory business, revenue is generated by attracting and retaining underwriting and other IB-business from corporations. This is achieved through issuance of optimistic research output as corporate managers are more favourable towards analysts and analysts' firms which support them. Lin and McNichols (1998) report that one of the factors that influences the selection of underwriters (by corporate issuers) depends on whether IB-analysts' reports portray a positive picture of the firms' operations. Hence, concern about attraction of IB-business motivates optimism among IB-analysts. Further, Ljungqvist et al. (2007, p.421) explains that:

> "Companies care about what the analyst has to say about their stocks and could take their investment banking business elsewhere if they are unhappy with the analyst's opinion. Thus sell-side analysts who work for integrated investment banking houses instead of research boutiques could come under implicit (or occasionally explicit) pressure to publish more favourable research about their employers' current or potential relationship clients to help boost investment banking fee revenue".

For analysts-firms with brokerage businesses, revenue is generated from commissions on stock trading. Analysts are also dis-incentivised from issuing negative reports on stocks covered as optimistic reports encourage trades and result in greater trading commission much more than pessimistic reports (Cowen et al. 2006). Primarily this is because trades can be initiated both by existing and potential shareholders (Fisch and Sale, 2002) if a 'buy' recommendation is issued, while pessimistic reports accompanied by 'sell' recommendations can only be acted upon by existing shareholders. Analysts personally benefit from generating IB-revenue and enhancing trade for their employers as their compensations and promotions were linked to their ability to provide such businesses.

2.4.3 Analysts' regulations

Analysts are an integral part of the securities industry. In the US[4], the Securities and Exchange Commission (SEC) is responsible for providing regulatory oversight. Part of this responsibility is delegated to several Self-Regulatory Organisations (SROs). Following the public outcry over the analysts' failure during the dot-com crisis and the alleged conflicts of interest that biased their research, several regulations were issued by the SEC and other SROs aimed at enhancing independence and objectivity of research analysts. These are briefly reviewed here and the empirical evidence on their effectiveness discussed[5].

[4] This study uses a US sample, therefore much of the discussion of the work of analysts is principally concerned with the US environment.

[5] A detailed review of all regulations affecting the work of analysts is beyond the scope of this discussion. The focus is mostly on recent legislations that has changed the work of analysts and as it relates to the conflict of interest that plagued their research. This provides necessary context for this study which is set in a post-regulatory environment and is particularly relevant for the second research objective of this research.

Regulation Fair Disclosure (Reg FD)

Reg FD was approved in August 2000 and became effective in October 2000 by the SEC. The rule sought to address the conflict of interest arising from analysts' dependency on corporate managers for information and eliminated the information advantage enjoyed by "favoured" analysts and institutional investors. To promote fairness and equity in the communication process between corporate managers and external constituents, it prohibited disclosure of material non-public information to analysts. In the case of inadvertent disclosures of such information in a private setting, corporate managers are required to make such information public by filing it with the SEC. Prior to Reg FD, selective disclosure of information was a means corporate managers used to reward optimistic and penalise pessimistic/critical analysts (Mayew 2008). Reg FD did not limit interactions with managers but rather the content and timeliness of such interactions.

The implication of this regulation is that some analysts are no longer to gain an informational advantage over others. It is expected that this will reduce analysts' dependence on corporate managers and the need to bias their research output to maintain close ties with them.

Several studies have examined the effectiveness of such regulation and the evidence is mixed. Chen and Matsumoto (2006) provide some support for the effectiveness of the legislation. However, Brown *et al*. (2015) report that analysts are still very much concerned about maintaining close ties with managers and this influences the conduct of their research. In addition to this evidence, Barker *et al*. (2012) examined why private interactions with managers are still a very important source of information for investors given the prohibition of disclosure of material information. The evidence from their study indicate that what is deemed relevant for assessing firm value includes confirmatory body language, tone and other factors and not necessarily discovery of new information. Soltes (2014) also provide support for the continuous importance of private meetings and close ties with between corporate managers and analysts. They report that access to managers is required for various other reasons beyond gaining informational advantage as addressed by Reg FD, including need to facilitate meeting between managers and investors (also referred to as "corporate access"). Westphal and

Clement (2008) also reveal that the benefit that accrue to analysts as a result of close ties with management and issuing research that favours corporate managerial interests goes beyond professional favours and include personal favours. Thus, though informational advantage may have been limited by Reg FD, evidence suggests that analysts still seek to maintain access the managers, there are still benefits to doing so beyond informational advantage and are still incentivised to issue optimistic research.

NASD Rule 2711 (Research Analysts and Research Reports) and NYSE Rule 472 (Communication with the public)

The National Association of Securities Dealers (NASD) – now Financial Industry Regulatory Authority (FINRA) and the New York Stock Exchange (NYSE) are self-regulatory organisations with some responsibility of issuing regulations to govern the conduct of operators in the securities industry in the USA, including analysts. In an attempt to sever ties between corporate finance and research departments within investment banks, NASD and NYSE issued rule 2711 and 472 in July 2002. These rules provided several measures directed at limiting the communication and influence of investment bank personnel on research personnel, required several disclosures aimed at promoting transparency and informing investors of any potential source of conflict and enable them make judgements on the quality of the research. These measures include (Guan *et al*. 2012):

- Disclosure of conflicts of interests in research reports
- Prohibits involvement of investment banking personnel influencing the content of research reports
- Prohibits analysts compensation from being related to IB transactions
- Prohibits the managers of covered companies from reviewing a research report before publication (except checking for factual accuracy)
- Research reports must explain the meaning of their stock ratings/recommendations
- Disclose whether any compensation is based on IB revenue, whether they hold a position as officer or director in the covered company and whether the company is a client of the analyst-firm.
- Disclose the percentage of recommendations that are “buy”, “hold” and “sell” categories.

The Global Analyst Research Settlement

The global settlement is an out of court settlement reached in April 2003 between the then New York Attorney General, SEC, NYSE, NASD and ten of the largest US investment banks. It was preceded by a series of investigations which revealed several cases of misconduct that violated independence of equity research. Thus, the main goal of the settlement was to enhance independence of equity research. The settlement consisted of a fine of over $1.4 billion on the banks and two research analysts[6] and involved several provisions which covered monetary relief, structural reforms, independent research and investor education (SEC, 2003). The settlement involved provision of monetary relief to certain investors, hurt by the misconduct of the banks. In total, the banks were to disburse over $875 million dollars to eligible investors. Structural reforms aimed at ensuring the physical separation of equity research and investment banking activity and limiting the flow of information between these two departments were also required. Measures include the prohibition of funding equity research with investment banking revenues. Investment banking personnel were prohibited from taking part in evaluating performance of equity research analysts or determining their compensation; research analysts were prohibited from participating in efforts to solicit investment banking business such as roadshows. For five years, each analyst-firm was required to provide, alongside its own research, research reports issued by independent research firms on the companies covered (SEC 2003).

In conclusion, these regulations have significantly changed the landscape of equity research. Given less funding for research which had been subsidised with investment banking fees in the past, there has been a decrease in both analysts' headcount and compensation at investment banks (Der Hovanesian and Borrus, 2005) resulting in talent exodus to buy-side firms, independent research firms and pure brokerage houses (Groysberg and Healy 2013). As the context for this study is post-regulation, this provides a context for understanding this study.

[6] The ten firms were Bear, Stearns & Co. Inc. (Bear Stearns), Credit Suisse First Boston LLC (CSFB), Goldman, Sachs & Co. (Goldman), Lehman Brothers Inc. (Lehman), J.P. Morgan Securities Inc. (J.P. Morgan), Merrill Lynch, Pierce, Fenner & Smith, Incorporated (Merrill Lynch), Morgan Stanley & Co. Incorporated (Morgan Stanley), Citigroup Global Markets Inc., f/k/a Salomon Smith Barney Inc. (SSB), UBS Warburg LLC (UBS Warburg), and U.S. Bancorp Piper Jaffray Inc. (Piper Jaffray) and the two analysts were Jack Grubman of Solomon Smith Barney and Henry Blodget of Merrill Lynch (SEC 2003).

The evidence from empirical research doesn't provide unequivocal evidence of the effectiveness of these regulations. If anything, much has been revealed that necessitates further investigation. Given regulatory focus on recommendations, there might have been a shift in strategy by analysts towards other output such as earnings' forecasts. For instance, Guan *et al*. (2012) found that while there has been a decline in optimistic recommendations, earnings' forecasts remained optimistic. This provides a motivation to investigate other elements of the reports beyond recommendation such as the written analysis. Moreover, the regulations have not addressed analysts' incentive to generate revenue for their employers through trading commissions which may motivate analysts' bias (Guan *et al*. 2012). This is also particularly important, if not more important (as reported in Cowen *et al*. 2006).

2.5 Chapter summary and conclusion

This chapter was aimed at providing background information on sell-side analysts, which are the object of this thesis. The rationale was to provide relevant information which serves as a context for understanding the literature and empirical analysis discussed in subsequent chapters of the thesis.

The chapter began by describing the role of sell-side analysts as information intermediaries. It further explains the impact of their activities on capital markets by improving informational efficiency and acting as external monitors of corporate managers. The activities that characterise information intermediation role of analysts was further described which consists of collecting, analysing and disseminating information to existing and potential investors. Further discussion centred on the analysts' reports as a medium through which analysts disseminate information to investors. The features of the reports were highlighted which include, timing and number of issuance, format and content of the reports and authorship. Of crucial relevance is that the reports vary significantly both in form and content. This raises an important question as to whether analysts' use of information in the reports and how they are presented is motivated by valuation or the need to sell a story about the companies. Also, existing literature has not extensively investigated the variation in the reports and the factors which may explain it; hence, the need for additional research.

Finally, the institutional setting and regulation which influences the work of sell-side analysts were also briefly reviewed. First distinction was made between IB and IND-analysts, which is relevant for this study. Second the conflicts of interest that influence the work of analysts were discussed. Third, the regulations enacted to address these conflicts were discussed. Focus was on the US setting, given the scope of this research.

Chapter 3: Analysts' use of accounting information

3.1 Introduction

The first research objective, which is to ascertain whether the extent of use of accounting information in analysts' reports vary with firm, analysts and report characteristics, is addressed in this and the next three chapters of the thesis. This chapter presents a review of the relevant empirical literature and begins by describing the literature search procedure in section 3.2. This is followed in section 3.3 by a review of the literature on the importance and use of information by sell-side analysts, with particular emphasis on accounting information. Section 3.4 discusses the gaps and re-states the research objective and related question. The suitability of content analysis in addressing the research question is discussed in section 3.5, while section 3.6 presents the theoretical framework relied on to formulate testable hypotheses. Finally, section 3.7 concludes the chapter.

3.2 Literature search procedures

The main search engine used during the course of the review process was 'Google Scholar', complemented by Business Source Premier. This was preferred to other databases for a number of reasons. First, the output of any search process includes both working papers available at SSRN or University web pages and other online repositories, books, theses, research articles and other relevant documents, it provides an extensive coverage of all scholarly work thus the number of relevant hits for any keyword search is higher. Moreover, for each search result, Google Scholar also provides several useful metrics such as number of citations and links to the papers citing the research article. These metrics have been useful in understanding the significance of research articles and have been used in the literature search process in locating sources of relevance for this study. For some highly relevant research papers, resulting from the search process, the list of associated citations was examined to identify other, more recent, relevant papers. Furthermore, the search process also involved inspecting the reference list of some papers to identify prior studies of relevance.

Keywords searched in this process included: *analysts* (and other variants such as investment analysts, sell-side analysts, financial analysts, security analysts[7]), *analysts' reports*, *interviews*, *protocol analysis, content analysis*, *relevance, accounting information etc*. These keywords were often used jointly during the search process.

[7] The search for the keyword analysts often returned results from other disciplines such as medicine and information technology. Hence, the need to qualify the term.

3.3 Use of accounting information by sell-side analysts

This section reviews the literature on the use or perceived usefulness of accounting information to analysts. A variety of research methods have been used to explore this topic empirically, such as interviews, questionnaires, protocol analysis and content analysis, with each method providing a different perspective on the subject.

Consistent with the broad definition of accounting information adopted for this study as explained briefly in Chapter 1 and in detail in Chapter 5, the review comprises studies which explore the use and perceived importance of types of accounting information, sources of such information and accounting based valuation models. For the purpose of the review, preference is given to financial analysts[8] rather than investors or other capital market participants, given the focus of this study[9].

3.3.1 Types of accounting information

Financial analysts rely on various types of information to derive earnings, price forecasts and make investment recommendations. The extent of use of accounting information or its perceived usefulness is inferred empirically through analysts' ratings of accounting information among other types of information (using interviews/questionnaires), observation of analysts' use of accounting information for equity valuation (using protocol analysis, experiments or participant observation) or examination of the number of references to accounting information in analysts' reports (using content analysis). The key findings across these studies are reviewed below.

In the US, Chugh and Meador (1984) report that expected changes in EPS, industry prospects and expected ROE are considered most important variables over the

[8]The term financial analysts is used here rather than sell-side analysts as several studies survey broad range of finance professionals which they often refer to as financial analysts. Hence, to ensure that the discussions are accurately reflecting the authors' views, the term has been retained. However, where distinctions are made between sell-side analysts and others, the review focused on discussing the evidence as it pertains to sell-side analysts, given the scope of this research.

[9] Cascino *et al*. 2013 provide a recent review of the literature on the use of accounting information by investors and debt providers in general.

long and short term for equity valuation and qualitative information about strategy, market position and quality of management is used to validate quantitative measures of performance. Similarly, Block (1999) reported that analysts ranked earnings as the most important input for valuation, followed by cash flows, while dividends and book value were considered least important. Results based on studies of analysts' reports are supportive of these findings. For instance, for a sample of 976 reports, Govindarajan (1980) found more references to earnings than cash flows. Similarly Previts *et al*. (1994) and Rogers and Grant (1997) document that financial and operating data were most recurrent themes in analysts' reports, with particular emphasis on financial performance related metrics such as EPS.

Results from studies outside the US (which are mostly UK based) are largely consistent with the findings from US studies. Coleman and Eccles (1997) found that cost, segment performance data, earnings and cash flows are highly ranked by analysts in conjunction with qualitative information about market share and market growth. Similarly, Barker (1999) reported that analysts prefer performance ratios based on the income statement over balance sheet related ratios, while Breton and Taffler (2001) documented that profitability related information was most prevalent in analysts' reports. These findings lend support to the theoretical models and numerous empirical findings in the value relevance literature which have long established and documented the relevance of earnings for equity valuation.

Given the pre-eminence of earnings, some studies focus only on understanding the perception of earnings' quality by analysts through assessment of their reports (e.g. Bricker *et al*. 1995) and follow-up interviews (Barker and Imam, 2008). The studies reveal that the concept of earnings quality is interpreted by analysts in light of both accounting and non-accounting information. With regard to accounting information, analysts consider the sources of income (whether income is from core operations or extraordinary, non-recurring activities), the use of conservative accounting policies and the impact of accruals. The predictability and sustainability of earnings are considered important attributes as well as the association between earnings and operating cash flow. Non-accounting based considerations involve assessment of the market conditions and the quality of

management (Barker and Imam, 2008). These studies are focused on the perception of earnings' quality rather than analysts' use of accounting information in general, nevertheless they provide useful insight into how analysts use information. First, analysts are concerned with earnings quality and do judge the quality of earnings using other accounting and non-accounting information. Second, non-accounting information complements accounting information rather than substitutes it.

Furthermore, prior research has largely ignored the determinants of analysts' use of accounting information. Of the explanatory variables studied to date, the most recurrent and consistent determinant of analysts' use of accounting information is *industry*. For instance, Coleman and Eccles (1997) reported that all financial analysts following pharmaceutical firms ranked information about product development, R&D investment, intellectual property and R&D productivity more highly than other analysts. Moreover, analysts following regulated utility firms such as water service companies were more concerned with environmental compliance data than other analysts. In a similar vein, Matsumoto *et al*. (1995) reported significant differences between analysts following firms in the retail and manufacturing sector. Inventory turnover, receivables turnover and gross margins were the most preferred information types for retail firms while R&D expense/sales, price/sales and price/book values were most useful for analysts following manufacturing firms. Industry effect was also reported in Glaum and Friedrich (2006) as they found that analysts following companies within the telecommunication industry generally preferred sales and cash-flow related information over earnings. Abdolmohammadi *et al*. (2006) also showed that analysts following firms in intangible intensive industries use less financial information than analysts which follow firms in industries that rely more on tangible assets.

Beyond industry, Nielsen (2007) revealed a strong positive association between analysts' employers' trading volume (a proxy for size of the analyst-firm) and the extent of use of various information types. Coram *et al*. (2011) also found that the extent of use of non-financial information by analysts is dependent on the financial performance of the firm. Specifically, they found that non-financial measures were more important to analysts when financial performance is positive.

Recommendation types have also been examined as a possible explanatory factor. While Govindarajan (1980) found no significant difference in the use of earnings over cash flow across recommendation types, Breton and Taffler (2001) found the relevance of information types to differ across recommendation types. Positive references to profitability and positive and neutral references to management and strategy were found to be positively associated with 'buy' recommendations as opposed to 'hold' or 'sell' recommendations.

In general, the evidence from these studies reveal that analysts value and use accounting information, particularly earnings, profitability ratios and related financial performance data. Most of the studies focused on information usage in general with only few studies examining variation in usage and with limited variables such as industry classification.

3.3.2 Sources of (accounting information) used by analysts

In the US, Chugh and Meador (1984) and Epstein and Palepu (1999) report that source of information preferred by analysts is interactions with company personnel through private contact or analysts' meetings, followed by annual reports which were generally perceived as important with the exception of the balance sheet. Studies based on protocol analysis further reveal that the income statement are the single most important source of information within the financial statement (Biggs 1984, Anderson 1988).

Elsewhere, some studies support the US evidence of analysts' preference for private interactions with company management over financial statements and related reports (Pike *et al*. 1993, Barker 1998 and Glaum and Friedrich 2006). However, other studies report that analysts' highly rank annual reports and interim statements over private meetings (Arnold and Moizer, 1984; Moizer and Arnold 1984; Vergossen 1993).Within the financial statements, the income statement is the most useful source of information, followed in some cases by segment reports and to a lesser extent, the cash flow statement (Barker, 2001; Glaum and Friedrich, 2006) while management notes and the balance sheet are least useful. Interview evidence, which enables further probing, offers some explanations for these findings. For instance, Barker (1998) found that financial analysts prefer face-to-face interactions as a source of information for various

reasons, including the opportunity to ask follow-up questions. Further, segment performance data provide additional "insight into company performance" (Coleman and Eccles, 1997). Moreover, the less reliance on management reports (despite containing forward-looking information) is due to a lack of content credibility as analysts believe managers are likely to present a positive view of their firm (Campbell and Slack, 2008).

For content-analytic studies, the approach to investigating sources of information is based on tracing information contained in analysts' reports to financial documents such as the annual reports. Rogers and Grant (1997) provide the first example of this type. They attempt to identify the extent to which the annual reports were used as a source of information for analysts by tracing the information cited in their reports to the companies' annual reports. A similar approach was adopted by Fogarty and Rogers (2005) and Abdolmohammadi *et al.* (2006). The former considered several information sources beyond the annual reports including press releases and Wall Street Journal articles while the latter examined only documents filed with the Securities and Exchange Commission (SEC). The consistent finding across these studies is that the annual report is a principal source of information for analysts as over half the information disclosed in the reports could be traced to the annual reports (Rogers and Grant 1997, Fogarty and Rogers, 2005) or the SEC filings (Abdolmohammadi *et al.*, 2006[10]). However, a caveat to note when interpreting the findings from these studies is that types of information need to be distinguished from sources of information[11]. Moreover, the successful tracing of information from the analysts' reports to annual reports does not necessarily imply that annual reports were used as a source for the particular information item as "accounting information may be an important feature of both personal contact and of organised visits and presentations" (Barker 1998, p.12) and therefore annual reports may only serve to confirm prior expectations. Moreover, Barker (1998) found that the low ranking of the annual reports by analysts (in their study) was mainly due to their timeliness as an information source rather than the irrelevance of accounting information as such. In summary, these studies suggest that the principal source of information

[10] This study also distinguishes between financial and non-financial information and documents that more financial information is sourced from the SEC filings than non-financial information.

[11] An observation previously addressed in McInnes *et al.* (2007).

for analysts is the firm itself. The findings are split between preference for personal contact with firm personnel through one-to-one meetings, analysts' presentations and company visits and corporate reports in the form of prelims, interims statements or annual report.

3.3.3 Types of accounting valuation models

Barker (2001, p.47) defined valuation models as "a means by which accounting information is related to share prices". They assume that the choice of a valuation model is indicative of the relevance placed on the underlying accounting information used in such models. Thus, inferences about the use of accounting information may also be based on the choices made about valuation models used.

In the US, Block (1999) report that most analysts do not rely on present value techniques. Similarly, Bradshaw (2002) found that recommendations are generally based on PE ratios. Moreover, research in other countries such as the UK reveal that the single most important valuation model for financial analysts is the price-earnings (PE) model (Pike *et al*. 1993; Arnold and Moizer, 1984; Moizer and Arnold 1984; Barker 1999, 2001). However, while early studies document the lack of use of discounted cash flow methods (DCF) and other sophisticated valuation models due to low reliability in estimating inputs (Barker 1999), more recent evidence suggest that DCF models are becoming increasingly popular (Glaum and Friedrich, 2006 and Imam *et al.*, 2008). It was found that concerns about increased scrutiny over financial analysts' activities following the dot.com crisis motivated the use of more sophisticated methods of analysis such as DCF models (Glaum and Friedrich, 2006). Further, Imam *et al*. (2008) document that the choice of valuation models is influenced by considerations of the company's financial statement quality, dividend-paying ability and understandability of the models. Moreover, Demirakos *et al*. (2004) found that analysts' use of valuation models is contingent on the industry of the firm analysed. Specifically, they found that a multi-period cash-based valuation model is likely to be used in high growth sectors as opposed to single-period comparatives. Accrual/earnings-based models are more commonly used in stable industries.

3.4 Research gaps and statement of objectives

The above review revealed that accounting information is important to sell-side analysts and the findings are generally consistent across topics. The importance of accounting information, particularly earnings-related information, is reflected in the income statements being the most useful source of information within the financial statements and the dominance of PE multiples as the preferred valuation model. Moreover, the findings are also consistent across the different research methods reviewed.

The review also revealed that much of the research attention has been focused on describing what and how types of accounting information is used by analysts with little attention given to the study of variation in the use of accounting information and the determinants of such variation. Therefore, there is need to extend existing literature in this direction. The objective of this part of the thesis is to address this gap by investigating the variation in the use of accounting information by sell-side analysts and identify the factors which explain such variation, through examination of the content of their equity research reports. The study seeks to understand whether company, analysts and report characteristics influence the extent to which analysts use accounting information in their reports.

Such a research endeavour is important for several reasons: First, it improves our understanding of analysts' decision making process by uncovering the determinants of their choices of input (information types) to their decision making. Second, knowledge of the factors which influence the extent of use of accounting information is important for policy makers as it enables them tailor responses to issues of irrelevance of accounting information accordingly. Third, knowledge of the determinants of use of accounting information is also useful for corporate managers. ICAEW (2009, p.7) acknowledged that "In deciding what to disclose, it is sensible for individual businesses to engage with their stakeholders and other users of their reports. In this way, they can judge what works and what does not, and where additional disclosures may be needed."

The following research question is addressed:

> What factors explain the variation in the use of accounting information by sell-side analysts?

This is investigated through examination of the content of analysts' reports. In addressing this question, the study purports to extend prior literature in several ways. First, accounting information is clearly distinguished from other financial information categories, unlike previous studies which generally classify accounting information into wider categories such as financial information. Second, sub-categories of accounting information are examined, which provides additional insight into uses of different types of accounting information. Third, this study distinguishes between historical and forward-looking accounting information, which enhances understanding of how accounting information is used. Fourth, a broad range of factors which explain the use of accounting information are examined. Fifth, the difference in the use of accounting information between analysts employed by investment banks and independent research firms are examined, which enables further investigation of the role of incentives in shaping analysts' decision making.

The next section justifies the use of content analysis over other research methods as have been previously used in this literature.

3.5 Suitability of content analysis

As previously reviewed in section 3.3, several methods have been used to investigate the relevance of accounting information by analysts. Each method provides different insights into how analysts use accounting information and therefore provides complementary evidence to studies based on other methods. For instance, the use of questionnaires allows a large sample to be examined which enhances external validity, while interviews provide the opportunity for the researcher to deepen understanding. The study by Barker (1999) is illustrative of this point; evidence from questionnaires revealed that analysts tend to highly rate unsophisticated valuation models such as PE multiples over more sophisticated models and accounting information over qualitative information about strategy. Further probing during interview stages revealed that the choices were mostly driven by consideration of reliability of the information type. It was found that "the value-relevance (and therefore the usefulness) of information varies according to the reliability with which the information can be forecasted..." and accounting information was found to be "intrinsically reliable", hence its relevance (Barker 1999, p.204). Future cash flows are also deemed too uncertain to allow value to be derived reliably when used in valuation models, hence the preference for single-period multiples.

Similarly, Barker (1998) observed that the importance of discussion with management (as a source of information) is due to several factors, such as the timeliness of the information and possibility to ask follow-up questions which provide avenue for researchers to obtain competitive advantage over peers. Moreover, accounting information was found to be relevant for two reasons (Barker, 2001): i) as a source of new information to the market through preliminary reports, and ii) as a source of reference for confirming or refuting previous expectations (about future financial performance). Not surprisingly, the evidence from participant observation is that analysts focus on earnings only at the initial stages of analysis (Barker, 2000). An advantage of the survey approach is that it directly investigates analysts' opinion and provides the opportunity to understand the rationale for analysts' choices. Inferences from survey studies are based on what analysts "say" they use or they find relevant. However, the validity of the results from surveys is contingent to a large extent on analysts' self-insight and recall abilities.

Furthermore, the use of protocol analysis offers the opportunity to assess analysts' use of information across different stages of the investment process. For instance, Biggs (1984) examined both the information gathering and evaluation stages of the analysis process, while Bouwman *et al.* (1995) split the analysts' activities into four stages as follows: familiarising, exploring, reasoning and scanning. The consistent finding is that accounting information and annual reports play a significant role as a source of information only in the early stages of the investment process. This was also reported in Barker (2000) based on participant observation.

Despite the benefit of providing incremental information about stages of the valuation process in which accounting information is most relevant, the experimental settings of protocol-analytic studies limit the generalisability of the findings. Inference can only be made based on the case presented to research participants and the information supplied, hence "the primary sample variation is the individual analysts rather than the analysed companies" (Bradshaw 2002, p.29). For instance, Biggs (1984) examined the information search strategy of 11 financial analysts. However, only financial statement information was presented for consideration to the analysts, so analysis and interpretation could only be made on that basis. In real life settings, financial analysts have a wide range of information available and therefore findings in an experimental setting which mirrors only part of the real life setting may not be very representative of actual practice, thus limiting the generalisability of these studies. Moreover, inference can only be made based on what participants actually document. It has been acknowledged that the requirement to verbalise thought processes for protocol analysis might influence participants' actual behaviour (Anderson, 1988). There may be difficulty in verbalising and analysing information concurrently. For instance, Bouwman *et al.* (1995) eliminated two analysts from their sample who admitted the difficulty in thinking and speaking aloud concurrently. Coram *et al.* (2011) further acknowledged that another limitation of protocol-analytic studies is that analysis can only be based on what is verbalised which may not correspond to what was actually considered or used in decision making. Hence, response bias might influence the results obtained from these studies. These limitations are exacerbated by the limited sample size of these studies.

Of these methods, content analysis is deemed most suitable to address the research objective stated above. Content analysis as a method of enquiry has been widely used in various fields of study such as politics, nursing and linguistics and has been the principal means of investigation of corporate voluntary disclosure within the accounting literature (Abhayawansa, 2011). Specifically, it is used to investigate various topics in the area of accounting, including corporate social reporting disclosure (see for instance, Gray *et al.*, 1995; Unerman, 2000), intellectual capital reporting (Beattie and Thomson, 2007); accounting narratives in general (Beattie *et al.* 2004) and disclosures by financial analysts (Barker and Imam, 2008; Abhayawansa and Guthrie, 2012). In relation to analysts' reports, content analysis has been successfully used to investigate several aspects of analysts' work process, such as the source of analysts' information (Roger and Grant, 1997; Fogarty and Rogers 2005; Abdolmohammadi *et al.* 2006); the relevance of different types of information (Breton and Taffler, 2001); variation in the use of financial information by analysts across industries (Abdolmohammadi *et al.*, 2006); the use of intellectual capital and non-financial information by analysts and their determinants (Flostrand 2006; Flostrand and Strom, 2006; Garcia-Meca and Martinez, 2007; Abhayawansa 2010); the tone of analysts' reports and their information content (Twedt and Rees, 2012). There are several reasons why content analysis presents itself as the most suitable technique for investigating variation in the use of accounting information and the determinants of such variation.

First and most important for this research, content analysis method enables the variability in the use of accounting information across companies and analysts to be examined. Each report is written by an analyst(s) at a specific brokerage or research firm for a given company in a definite time period. Therefore, by examining the content and use of information across a variety of reports, the variability across period, analysts or brokerage house and the company characteristics can be investigated. This is rarely achievable through the use of other approaches such as interviews or questionnaires. Analysts follow and comment on several companies, however research using survey methods focuses on analysts' approach to valuation and use of information in general rather than eliciting their views with regards to a specific company with its unique features. Researchers rarely investigate the influence of contextual factors such as firm

size, etc. The study by Orens and Lybaert (2010) was an exception as the authors investigated the effect of several factors on analysts' use of non-financial information. The research relied on questionnaires as a method of data collection and involved analysts ranking their use of non-financial information on a five-point Likert scale. The scores derived from the survey instrument were regressed on the average leverage and stock volatility of the firms followed by each analyst. The authors assumed that analysts' rating of information types is an indication of the average use of information across the different companies they follow. However, this is an approximation which doesn't account for the variability in usage of information. Content analysis method overcomes the need for such approximation as each report is unique for each company.

Second, the use of content analysis overcomes the problems of recall and other self-reporting biases that characterise survey type methods (Hassan 2005). The issue of recall bias is particularly relevant for the proposed research due to the investigation of usage of accounting information across different company characteristics. Hence, to provide valid responses in a survey setting, analysts would need to recall their use of information across different companies and the characteristics of such companies.

Third, as noted by Breton and Taffler (2001, p.93), the use of content analysis in relation to analysts' reports is also advantageous due to the ability to capture the relevance of different information types used by analysts. This is possible due to the nature of the content analysis method in which frequency of occurrence of themes is considered an indication of its relevance.

A final rationale for the preference for content analysis is due to the advantage of a real life analysis. Analysts' reports used in this study are written in real life settings by analysts to investors. Hence, it allows the investigation of what analysts do in practice rather than reliance on what they say they do.

Moreover, this method is unobtrusive in nature, which allows investigation of the practice of sell-side analysts without the problems of accessing these groups, which is often very difficult to obtain. Most survey methods, where access has been possible, have very low response rate and small sample size which limits the generalisability of the findings. In contrast, analysts' reports are made available

to the public[12] and the only limit to the number of reports that can be analysed is the laborious nature of the manual content analysis process.

Against these advantages are several limitations, such as concerns over reliability issues due to the somewhat subjective nature of the content analysis process. However, with well-specified coding rules and appropriate research design, these issues can be overcome to a certain extent. Furthermore, with particular reference to analysts, it has been argued that it is likely that not all the information used by the analysts is disclosed in the reports (Schipper, 1991).

[12] However, access to the financial databases where these are stored such as Thomson One is based on payment of subscription fees.

3.6 Theoretical framework

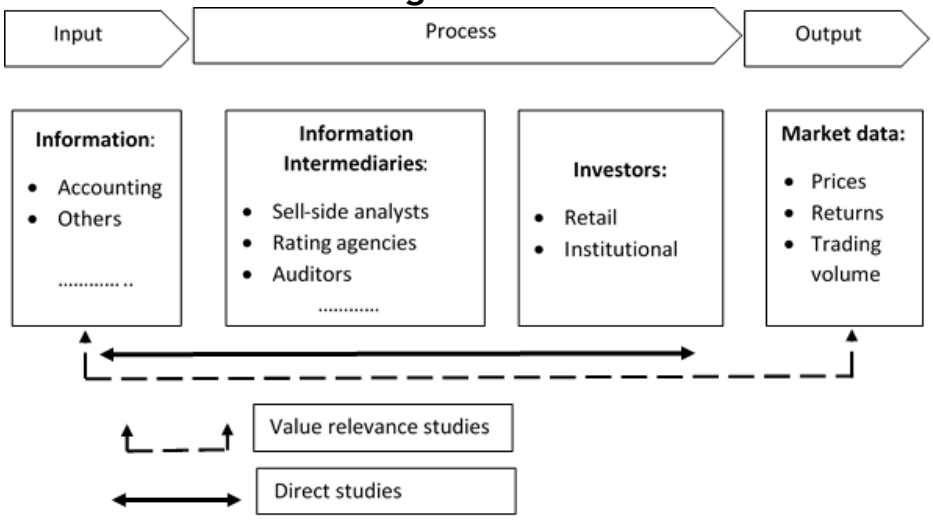
A review of the literature on analysts' information usage highlights a pervasive lack of theoretical framework in the design of research (Beyer *et al.*, 2010).

Why does the use of accounting information in analysts' reports vary? To address this research question, it is imperative to first identify the rationale for the use of accounting information by sell-side analysts. In this study, a valuation perspective is adopted as the principal reason why information is used by analysts. In principle, analysts' main activity involves the assessment of company performance and prospects, on the basis of which companies are valued and investment decisions recommended (Schipper, 1991; Healy and Palepu, 2001). Hence, analysts use information relevant for valuing firms and making investment recommendations. This is consistent with the CFA institute statement of practice which advocates that analysts' reports include information used to arrive at investment recommendation (CFA, 2010).

The usefulness of accounting information for equity valuation has been the focus of value relevance studies, which is relied on to provide a framework for understanding the role of accounting information in valuation by analysts. Barth *et al*. (2001) define value relevance of accounting information based on its association with prices or returns. Value relevance studies may be classified into two main categories: association studies and information content or event studies, (Holthausen and Watts, 2001). Association studies make inferences about the usefulness of accounting information based on the strength and statistical significance of the association between accounting information and prices or returns (over long horizons). Thus, relevance is considered as the ability of accounting data to summarise information that influences prices or the ability of accounting information to measure firm value (Francis and Schipper 1999). On the other hand, event studies make inferences about the relevance of accounting information based on price reaction to its release (over short windows). In this case, relevance is measured as the ability of accounting information to provide "news" that alters investors' prior expectations (Francis and Schipper, 1999). In each of these cases, actual usage of accounting information and users' preferences are not measured but inferred based on market values and returns. Nevertheless, some studies which examine the usefulness of accounting

information to diverse users such as analysts have cited value relevance literature in interpreting findings from their study (e.g. Barker and Imam, 2008). Perhaps this is a reflection of the similarity between the value relevance studies and direct studies which investigate the actual usage of accounting information. Figure 2 provides a simple illustration of the investment decision process and highlights the link between both streams of research:

Figure 2: Comparing direct and market-based approaches to relevance of accounting information



As illustrated, the input to the investment decision process is information. Information is obtained from various sources and includes both accounting and non-accounting information. The investment decision process involves the use of such information to assess firms' performance and prospects, on the basis of which investment decisions are made. Investment decisions in turn influence market prices, returns and trading volumes, which represents the output in Figure 2.

Information collation and analysis may be undertaken directly by investors or they may rely on information intermediaries such as analysts, who recommend investment decisions. Analysts are relied upon by investors to interpret existing public information and provide new private information (Chen *et al.*, 2010). The output of analysts' decision processes influence investors' trading behaviours, which in turn influences market variables. Empirical evidence reveals that analysts, through their intermediation, improve market reaction to new information, thereby improving market efficiency (Wommack, 1996, Jagedeesh *et*

al., 2004 and Asquith *et al.*, 2005). This implies that analysts play a significant role in ensuring that information is priced and therefore value relevant.

On one hand, value relevance studies consider only the relationship between the input (mostly accounting information) and output (market values) and ignore the decision making process by the intermediaries and investors (Barker, 2001). Thus, while providing evidence of stock market reaction to earnings announcements and the association between accounting information and value, it is silent about the underlying process which ensures that information is reflected in stock prices (Barker, 1998). On the other hand, direct studies examine use of accounting information from the perspectives of capital market participants and they enhance our understanding of the price formation process and the role of intermediaries such as analysts (Brown, 1993). These studies consider the usefulness of different information types and sources to investors and information intermediaries and how such information is used. This thesis contributes to this stream of research by examining the factors which explain the variation in the use of accounting information by analysts.

Given the similarity between direct and the market based approach to studying the usefulness of accounting information illustrated above, this research relies on the value relevance studies to provide a basic theoretical understanding of the use of accounting information in equity valuation, which is useful in formulating testable hypotheses. Ohlson (1995) is often cited as providing a theoretical model for value relevance studies (Easton 1999, Barth *et al.* 2001). He developed a model that expresses market value as a linear function of book value and the present value of expected future abnormal earnings. The underlying theory is the classic valuation theory in finance which expresses a firms' equity value as the present value of its future dividends or free cash flow to equity (Biesland, 2009). Beginning from this framework, he derived a valuation model which replaces future dividends with book value and contemporaneous earnings, based on several assumptions such as the clean surplus relation and linear information dynamics. Following from this, several studies indicate that the value relevance of accounting information varies across different firm characteristics (e.g. Hayn 1995, Burgstahler and Dichev 1997, Barth *et al.*, 1998) and these characteristics

are used as explanatory variables for this study. Chapter 4 discusses the rationale and formulates testable hypotheses.

3.7 Chapter summary

The purpose of this chapter was to review the literature on the use and relevance of accounting information to analysts and to identify the limitations and gaps in prior research as well as introduce the research objective of this study and theoretical framework.

The chapter commenced with a discussion of the literature search methods and was followed by a review of the literature which uses direct methods to investigate the use and relevance of accounting information to analysts. The review revealed that accounting information is useful to analysts, particularly financial performance-related information such as earnings. It was further observed that there is a dearth of research which investigates the factors that influence analysts' use of accounting information and most analysis of the content of analysts' reports has categorised accounting information into a "financial information" category such that the overall findings with regards to accounting information is inconclusive. Thus, the research objective of this study emerged from this review and involves the investigation of the variation in the extent of use of accounting information across firms, analysts and report characteristics with the aim of identifying the factors which explain such variation. Given prior research have adopted a variety of methods to investigate analysts' use of accounting information, this chapter also reviewed the benefits and limitations of the various methods used and justified the choice and suitability of content analysis as the method of enquiry for addressing the research question. Finally, the value relevance literature is introduced as providing a useful framework which is relied on to formulate hypotheses for this research and interpret findings.

Chapter 4: Hypotheses development I

4.1 Introduction

The purpose of this chapter is to develop and discuss testable hypotheses in relation to the previously stated research objective and question. As discussed in Chapter 3, a valuation perspective is adopted as the main rational for analysts' use of accounting information in their reports, hence factors that influence usefulness of accounting information for valuation purposes are the main focus. The value-relevance literature is relied on to provide a framework for the selection of explanatory variables used for this research. The principal motivation is that analysts are generally concerned with company valuation. Hence, the relevance of accounting information for such purposes is in line with the value-relevance literature and the theoretical models on which empirical analysis is based. Flostrand and Strom (2006, p.580) explains that "information has valuation relevance if it is used by analysts in the valuation process" and information contained in analysts' reports is indicative of its valuation relevance. However, there are several reasons why the evidence and arguments advanced in explaining the variation of the use of accounting information across firm characteristics may be limited in this study.

First, value relevance studies consider only the relationship between the input (mostly accounting information) and output (market values) and ignore the decision making process by intermediaries and investors (Barker, 2001). Thus, while providing evidence of stock market reaction to earnings announcements and the association between accounting information and value, it is silent about the underlying process which ensures that information is reflected in stock prices (Barker, 1998). Behavioural factors that may influence analysts and investors uses of accounting information are not modelled. Moreover, non-rational factors may also play a role. Abhayawansa et al. (2015) noted that analysts do not act in an objective and rational manner as suggested by neoclassical theory and use information strategically (see also Beunza and Garud (2007) and Abhayawansa and Guthrie (2012)).

Second, value relevance literature is principally focused on earnings and to some extent book-value of equity and it is not clear how and whether the findings and

arguments in relation to earnings and book-values extend to other types of accounting information which is investigated in this research. Hence, the factors previously documented as influencing the price-earnings and return-earnings relationship in the value relevance literature may differ for other types of accounting information.

Third, the concept of relevance is differently construed. In the value relevance literature, the relevance of accounting information is based on the strength of statistical association between market data (such as prices and returns) and accounting information. However, this thesis examines the extent with which accounting information is discussed in analysts' reports. Hence, it focuses on analysts and their reports rather than the market. Prior research suggest that usefulness of accounting information is not synonymous with 'price-sensitivity' as generally perceived in the value relevance literature. Information not value-relevant still retain usefulness to users such as fund managers (Barker et al. 2012) and sell-side analysts (Abhayawansa et al. 2015). As noted in Abhayawansa et al. (2015), information (such as intellectual capital information) is used at different stages in the valuation process such as estimating/forecasting earnings and cash flows, setting an appropriate discount rate in the DCF framework, communicating with clients or adjusting the target prices. Hence, the concept of relevance of accounting information is construed rather narrowly in valuation theory. Thus, factors that influence analysts' choice of information may in part differ from those suggested by valuation theory. This research provides an empirical test of the extent to which valuation theory explains analysts' use of accounting information.

Fourth, analysts' reports are not "records of their decision processes" (Breton and Taffler 2001, p.99) and reports "do not necessarily include all information used to arrive at a recommendation" (Rogers and Grant, 1997). To the extent that information is included in the reports for other reasons beyond their relevance for valuation, then a valuation perspective may be limited. A possible reason for citing information in reports could be for communication purposes. Analysts' reports are a means for communication, thus analysts might be more interested in communicating their opinions or "selling a story" to their audience rather than disclosing the information sets used in their valuation process. Consequently, whether valuation concerns influences analysts' use of accounting information and whether valuation theory explains analysts' use of accounting information is an empirical question.

This study examines the firm-specific variables suggested by valuation theory. However, other reasons may be advanced to motivate the selection of variables as discussed above. The focus on actual users rather than market statistics imply that behavioral factors may influence the selection of information by analysts. Analysts' cognitive abilities and incentives may play a role in their selection of information. Moreover, the focus on analysts' reports imply that there may be a mismatch between information used in the valuation process and information used for communicating with investors. Hence, in the selection of variables, the firm-specific attributes suggested by the value relevance literature were further complemented with variables which capture the analysts-specific attributes and the recommendations contained in their reports.

Section 4.2 outlines the hypotheses for firm-specific variables, section 4.3 for analyst-specific variables and section 4.4 for report-specific variables. Section 4.5 further explains the research design for testing the hypotheses and the selection of proxies for the variables used in the analysis. Finally, section 4.6 provides a summary of the chapter.

4.2 Company- specific features

As stated in the previous chapter, the main theoretical framework which guides the hypotheses formulation for this study is the theoretical modelling of accounting and firm value within the value relevance literature. Within this literature, the relevance of accounting earnings for valuation is derived from the linear information dynamics assumption of an autoregressive time-series property of earnings (Lo and Lys, 2000, p.348). This implies that earnings' role in equity valuation is dependent on the persistence and sustainability of current earnings. Earnings persistence and sustainability is also referred to as the quality of earnings in Dechow *et al.*, (2010), and is consistent with analysts' interpretation of earnings' quality in Barker and Imam (2008).

Given a valuation perspective, it is posited that analysts use of accounting information is based on its relevance for their valuation and forecasting purposes[13]

[13] This stance is taken based on the statement in CFA institute's professional ethics guidance which suggests that information relevant for arriving at recommendations and price and earnings forecast be included in analysts' reports. There are several reasons why analysts could use accounting information in their reports beyond its relevance for their recommendations (e.g.

and relevance is influenced by analysts' perception of the quality of accounting information (Barker, 1999) which is determined by several firm characteristics (Dechow *et al.*,2010). According to Barker (1999, p.204), "the value-relevance (or usefulness) of information varies according to the reliability with which the information can be forecasted." Factors that impact on earnings' persistence and reliability are therefore predicted to influence its relevance to analysts and thus the explain variation in the extent of use of accounting information in their reports. Several studies within the value relevance literature have examined a number of variables that influence the relevance of accounting information for equity valuation. These studies are relied on to justify the firm-specific variables examined in this study.

Value relevance literature has mostly focused on a few accounting items primarily earnings and to a limited extent, book value of equity. Similarly most of the discussion about quality of accounting items focuses on *earnings'* quality. However, this study examines a broad range of accounting information (as explained in Chapter 1 and further defined in Chapter 5 of this thesis) and it is not clear how the findings and arguments in relation to earnings and book values extend to other accounting items. Variations in the quality and relevance of earnings may have a positive, negative or no impact on the use of other types of accounting information by analysts. Moreover, the impact may vary across different accounting items such that the overall effect is neither positive nor negative. Consequently, hypotheses are specified at the broader level of "accounting information" and are non-directional. However, detailed test will be conducted at lower level categories.

4.2.1 Financial performance

The level of financial performance is considered one of the most important variables that influence the relevance of earnings (Collins *et al*. 1997, 1999). The rationale being that losses or low levels of earnings are transitory in nature as shareholders have the option to liquidate their shares if they persist (Hayn 1995). Hence, losses are less indicative of future prospects for a firm than profits. Using

---

to satisfy their audience's information needs, to present a positive view of the firm, etc). However, the hypotheses are based on the CFA's statement of relevance as an a priori expectation which may be confirmed or refuted.

a sample of over 9000 firms from 1962-1990, Hayn (1995) found that the coefficient of earnings in the returns-earnings regression for a loss only sample is significantly less than those of profit only sample. Similar results were observed for low levels of financial performance in general, which they attribute to the option available for shareholders to liquidate their shares. Relatedly, Burgstahler and Dichev (1997) argue that the role of earnings and book value in equity valuation are complementary and differ based on the levels of earnings relative to book value. They show that at higher level of earnings, earnings are relatively more important determinant of value while book value is more important determinant of equity value at low levels of earnings (as the company's resources are more likely to be adapted for different purposes). This is supported by Collins *et al*. (1999) which document increased value relevance of book value for loss firms. They attribute this finding to the role of book value as a proxy for expected future normal earnings.

Darrough and Ye (2007) further suggests that several loss firms are also high growth firms or report losses due to significant investments with returns that are reflected in accounting records at a future date while being currently reflected in market prices. They argue that, given the nature of their operations and the fact that most investments are not recognised in traditional financial statements, both earnings and book value are likely not to be very important determinants of value. Thus, they conclude that both earnings and book value are insufficient to explain the value of loss firms in today's economy.

The first set of studies suggests that, for firms making losses, earnings are less useful in explaining value and the book values are more relevant. However, Darrough and Ye (2007) suggest that both earnings and book value are insufficient as other qualitative information beyond accounting information becomes increasingly important. This implies that analysts may rely less on accounting information in general and more on qualitative information for low levels of financial performance.

During the dotcom era, Penman (2003) observed that analysts largely ignored accounting information in favour of qualitative information such as number of clicks received by a dot.com company, with claims that the financial reporting was out of date and earnings were no longer relevant. At the wake of the crisis,

it appeared that the losses reported by the dotcom firms (and ignored by most analysts) were indicative of future problems. Hence, he argued that the weakened relation between accounting and value might not necessarily be a result of a faulty financial reporting system (as Darrough and Ye, 2007 suggest) but also of faulty pricing system which was speculative rather than based on fundamentals.

Poor financial performance also provides an incentive to engage in earnings and impression management (Clatworthy and Jones, 2006; Dechow *et al.*, 2010). Consequently, the reliability of earnings may be in question. Thus, analysts may rely less on accounting information due to concerns about its reliability (Barker, 1999).

Given the predicted impact of earnings is opposite for book value, it is not clear what the net impact is, nor is it clear what the impact on other types of accounting information is, which remains an empirical question. Thus it is hypothesised that:

H_1: The extent of use of accounting information by analysts is associated with firms' financial performance

4.2.2 Volatility of earnings

Theoretically, the role of earnings in equity valuation derives from its ability to proxy for future expected abnormal returns. The link between current and future earnings is weakened in the presence of highly volatile earnings (Dichev and Tang 2009). Such volatility may also be indicative of high information asymmetry. Consequently, highly volatile earnings are perceived as less predictable both by corporate managers (Graham *et al*. 2005) and analysts (Barker and Imam, 2008).

A fundamental step in analysts' valuation process is the forecasting of accounting earnings. Barker and Imam (2008) report that analysts are concerned about earnings' volatility and attempt to identify firms' core earnings when forecasting. Moreover, they consider that relevance of information varies with the reliability with which it can be forecasted (Barker, 1999). Hence, it is expected that volatile earnings are less relevant given their low predictability. Orens and Lybaert (2010) observed that earnings informativeness and relevance to analysts is negatively influenced by risk, proxied by volatility of stock return.

Thus, higher earnings volatility is expected to negatively influence use of accounting earnings. The effect on other accounting information still remains an empirical question. For instance, analysts may substitute earnings for less volatile income statement items such as revenues or may prefer to rely on use of non-accounting information. Hence, a non-directional hypothesis is formulated as follows:

H_2: The extent of use of accounting information by analysts is associated with variability of earnings.

4.2.3 Firm size

The effect of firm size on the value relevance of earnings has received significant attention in the literature. Collins and Kothari (1989) provided early empirical support for varying relevance of earnings due to firm size, where firm size was considered a proxy for information environment[14]. Larger firms have richer information environment characterised by increased disclosure, higher analysts' following, investor recognition and greater scrutiny. Consequently, they have lower information asymmetry which results in greater value relevance of earnings and book value (Bryan and Tiras, 2007)

Larger firms have also been found to possess certain characteristics that influence the informativeness of their financial statements: Larger firms are less likely to liquidate their operations and report losses (Hayn 1995); due to greater diversification, they also tend to be less risky. Dichev and Tang (2009) found that several proxies for size were correlated with low earnings volatility and conclude that larger firms have more predictable earnings. Taken together, the argument and evidence support a positive effect of firm size on value relevance of earnings. However, the effect on other types of accounting information remains an empirical question.

[14] Collins and Kothari (1989, p.145) defined information environment as including "all sources of information relevant to assessing firm value. It includes government reports on macroeconomic conditions, industry reports and trade association publications, firm-specific news in the financial press and reports issued by analysts and brokerage houses in addition to accounting reports, and vertical and intra-industry information transfers via sales and industry reports."

Prior study on the content of analysts' reports shows a positive relationship between analysts' use of non-financial information in the reports and the size of the company covered (Flostrand and Strom, 2006). This is attributed to the greater availability of information for larger firms. Whether this finding extends to accounting information is an empirical question. Hence, the following hypothesis is tested.

H_3: The extent of use of accounting information by analysts is associated with firm size.

4.2.4 Growth

Company's growth characteristics have been widely reported as influencing the accounting – value association. It is argued that accounting data is not as informative of firm performance for high-growth firms compared with low-growth firms. Growth opportunities are reflected in stock prices earlier than in historical accounting information and high growth firms "tend to have more of their value from future earnings" (Orpurt and Zhang, 2009). High growth firms also have lower earnings persistence (Dechow *et al*. 2010) and less predictable earnings. Prior empirical evidence supports a negative relationship between growth and the value relevance of both earnings and book value of equity (Frank, 2002; Ball and Shivakumar, 2008) and are frequently controlled for in the value relevance literature.

Flostrand and Strom (2006) found no association between analysts' citation of non-financial information in their reports and book-to-market ratios, which was used to proxy for growth opportunities. The results might be in part sensitive to the use of disclosure index or the proxy for growth opportunities used. Moreover, they focused on non-financial information. Given the limitation of accounting earnings in valuing high growth firms, analysts' may rely more on non-accounting information for such firms as was the case during the dotcom bubble (Penman 2003) or earnings might be substituted for other accounting information such as revenues. Thus the effect of growth characteristics on analysts' use of accounting information is empirically tested using the following hypotheses:

H_{4a}: The extent of use of accounting information by analysts is associated with sales growth
H_{4b}: The extent of use of accounting information by analysts is associated with market-to-book ratio

4.2.5 Financial leverage

Another variable of interest which has been found to influence the value relevance of book value and earnings is financial leverage. There are two prepositions regarding the influence of financial leverage. First, it is argued that the components of financial statements have differing roles as the income statement is primarily useful for valuation while the balance sheet provides information that is useful for debt contracting and particularly estimating a company's liquidation or adaptation value (Burgstahler and Dichev 1997, Barth *et al*. 1998, Collins *et al*. 1999). As, increased levels of leverage is associated with increased risk of bankruptcy, the liquidation value which is proxied by book value of equity becomes relatively a more important determinant of value than earnings. Based on this argument, Barth *et al*. (1998) estimate yearly cross-sectional regressions for ex-post bankrupt firms and found that as financial health deteriorates, the value relevance of earnings decreases while that of book value increases.

A second preposition advanced to explain the influence of financial leverage is that the quality and reliability of accounting earnings may be compromised for highly levered firms. It is argued that concerns over violating debt covenants may lead such firms to engage in earnings management (Dechow *et al*. 2010). Consistent with this argument, Ghosh and Moon (2010) observed a negative relationship between earnings quality and debt levels. Collectively, the evidence from prior value relevance studies is that high levels of debt increases the relevance of book value but decreases the relevance of earnings.

Studies that directly examine analysts' behaviour provide additional insight. For instance, Previts *et al*. (1994) found that analysts make extensive reference to liabilities in the report of firms which are highly levered. Similarly, Orens and Lybaert (2010) proposed and found that leverage is positively associated with analysts use of non-financial information, which they attribute to lower quality of accounting information for highly levered firms. Evidence from Barker and Imam

(2008) suggests that analysts are concerned about the quality of earnings. So, they may rely less on earnings when predicting value for highly levered firms given concerns over its quality. However, as sophisticated users, they may also be able to undo any earnings manipulation by managers, in which case, concern about earnings management would not be a determining factor. Moreover, most of these concerns relate principally to earnings and to a limited extent book value of equity and it is not clear what the effect on other types of accounting information is. Hence, the impact of leverage on the use of accounting information by analysts remains an empirical question tested by the following hypotheses:

H_5: The extent of use of accounting information by analysts is associated with leverage levels.

4.2.6 Industry

It has been argued that traditional financial accounting information is of little use to investors when valuing firms with high investment in intangibles. While market values reflect such investments, financial accounting only partly recognises such intangibles resulting in a weaker association between accounting information and value. Hitherto, empirical evidence has been mixed. Amir and Lev (1996) found that earnings, book values and cash flows are irrelevant on a stand-alone basis when valuing companies in the cellular telephone industry as they rely greatly on intangibles for value creation. Similarly, Collins *et al.* (1997) found decreased value relevance of earnings but increased relevance of book value for firms in industry with high investment in intangibles compared to firms in low-intangible industries. Contrary to these, Francis and Schipper (1999) found no difference in value relevance of earnings but increased relevance of book value for low-tech firms compared to high-tech firms.

Research on analysts' use of information also documents the relevance of industry affiliation and intangibles. For instance, Abdolmohammadi *et al.* (2006) found that analysts' reports for companies in industries which rely heavily on intangible assets contain less financial information than reports for companies in tangible-intensive industries. Flostrand and Strom (2006) failed to find a significant association between analysts' use of non-financial information in their reports and industry affiliation. Other studies including survey based papers also suggest that

analysts pay attention to the industry and this influences analysts' preference for different information (Govindarajan 1980, Matsumoto *et al*. 1995 and Coleman and Eccles 1997). Hence, it is expected that:

H_6: The extent of use of accounting information is lower for firms in the high-tech industry compared to firms in the low-tech industry

4.3 Analysts' characteristics

As previously argued in section 4.1, information cited in analysts' reports may be due to factors different from their relevance for valuation. Unlike the value relevance studies, direct methods such as analysis of the content of analysts' reports enable analysts' characteristics to be examined. This is particularly relevant given ICAEW (2009) conjecture that use of accounting information varies across users. Analysts are sophisticated users of accounting information (Schipper 1991) and are a particularly interesting group given their role in capital markets as discussed in Chapter 2.

Prior studies that examine the impact of analysts' characteristics on the valuation process have been mostly concerned with the output of the process such as recommendations, earnings' forecasts or price targets. There is little evidence on how analysts' characteristics influence their use of information. Prior behavioural accounting researchers have investigated the impact of the characteristics of the "decision-maker" on the use of information[15]. More recently, using surveys, Orens and Lybaert (2010) demonstrate that the use of non-financial information by analysts varies with their experience. Two analysts' characteristics previously found to influence their output are further tested as possible determinants of their use of accounting information in this study. These are discussed below.

4.3.1 Employer

Prior studies have investigated the characteristics of analysts' employers and how they influence analysts' research outputs such as their recommendations, earnings forecasts and price targets (Cowen *et al*. 2006, Jacob *et al*. 2008). This literature is reviewed in Chapter 7 of this thesis. As discussed in that chapter, the focus of these studies has been to understand how various business activities incentivise analysts to issue optimistic outputs. It has been observed that IB-analysts issue more optimistic recommendations and forecasts than IND-analysts. The effect of analysts' incentives to be optimistic (based on their employer's business activities) on their usage of information has been largely ignored.

[15] Libby and Lewis (1977) provide a review of these studies which rely on human information processing literature.

Moreover, as discussed in Chapter 2, investment banks and independent research firms differ on several dimensions beyond incentives to bias their research output, such as size, resources and information availability. For instance, Jacob *et al.* (2008) argued that investment banking analysts have access to superior information about the companies they cover as they provide services beyond just equity research, which enables closer relationship with corporate managers. Moreover, investment banks and brokerages are larger and have greater resources than independent research firms such that they could engage in detailed search for information beyond public available information.

Taken together, IB-analysts and IND-analysts firms differ significantly on several dimensions and it is expected that the effect of such difference is not limited to their output but also the input (information) used in their decision making. Hence,

H_7: The extent of use of accounting information differs between analysts employed by investment banks and brokerage firms and independent research firms.

4.3.2 Chartered Financial Analyst (CFA) qualification

The CFA qualification is a recognised and accredited structured training program which equips analysts with in-depth knowledge and skills in financial analysis. Though not mandatory, it is becoming increasingly popular among analysts and other finance professionals. A few studies have examined the benefit of the qualification by investigating the quality of analysts' research output such as accuracy of earnings' forecasts and performance of investment recommendations. For instance, De Franco and Zhou (2009) hypothesise and find that forecasts by CFA charter holders are timelier, bolder and less optimistic but inconclusive with regards to accuracy. Similarly, Kang *et al.* (2012) provide evidence of improved investment recommendation following CFA qualification. In view of the evidence that the CFA qualification impacts analysts' decision-making and the overall output of their research, this study extends the literature to examination of the impact of CFA qualification on analysts' use of accounting information (which is a significant component of the CFA curriculum).

H_8: The extent of use of accounting information differs between CFA and non-CFA analysts.

4.4 Report characteristics

In this section, the association between recommendation types and analysts' use of information is discussed.

Prior studies have investigated the association between analysts' recommendations and their use of information. Breton and Taffler (2001) examined the association between analysts' uses of different types of information (including accounting information) and their stock recommendations. They found that "buy" recommendations are associated with a higher proportion of references to management and strategy and positive references to the firms' profitability. Garcia-Meca and Martinez (2007) found a positive association between recommendation types and the extent of use of intellectual capital information in analysts' reports. While they offer no explanation for this relationship, Abhayawansa and Guthrie (2012) posit that the variation in the use of information across types of recommendations is indicative of impression management. They argue that intellectual capital related information is used to subdue pessimism associated with unfavourable recommendations and to improve the credibility of favourable recommendations. In a similar vein, Beunza and Garud (2007) reveal that analysts with different recommendations on Amazon.com during the dotcom bubble, used different accounting metrics in justifying their respective recommendations, with negative recommendations being justified by profit figures and more positive recommendations justified by revenue. The analysis was set in the context of the dotcom bubble which was characterised by relatively low profit for most e-commerce firms such as Amazon. However, it shows that different accounting information might be used strategically to justify different recommendations. Hence, the following hypothesis is tested:

H_9: The extent of use of accounting varies across recommendation types.

4.5 Testing the hypotheses

4.5.1 Statistical tests

The main analysis is aimed at examining the association between analysts' use of accounting information and company, analysts and report characteristics. OLS regression model has been used to examine association between content analysis scores and other variables as Flostrand and Strom (2006) and Garcia-Meca and Martinez (2007). However, for this research, a Tobit model is used, following Abraham and Cox (2007). The dependent variable in this study is a measure derived from content analysis of analysts' reports. The variable is censored in the sense that it cannot take on negative values. In such cases, OLS estimation techniques yield inconsistent parameters for the regression model (Long, 1997). A commonly used approach to overcome this limitation is estimation of a Tobit regression model, based on maximum likelihood. The estimation model is as follows:

$$ACC_{ij} = \alpha + \boldsymbol{\beta X_i} + e_i \qquad (1)$$

$$\text{With } y_{ij} = \begin{cases} ACC_{ij} & if\ ACC_{ij} > 0 \\ 0 & if\ ACC_{ij} \leq 0 \end{cases}$$

Where ACC_{ij} is the extent of use of accounting information or its sub-components in reports issued by analyst *i* for firm *j*. α is the constant term, β is the vector of coefficients on the independent variables and X_i is the vector of company, analyst and report characteristics which are presented in section 4.5.3 below and e_i is the error term.

Word count and Intangibles are control variables in the analysis.

4.5.2 Variables

The dependent variable used in this study (ACC) for testing the hypotheses is the extent of use of accounting information. This measure is derived from manual content analysis of analysts' reports. Chapter 5 explains the content analysis process in detail including categorisation scheme, coding units and categories of accounting information. The extent of use of accounting information (or its sub-categories) is derived as follows:

$$ACC = \frac{Number\ of\ words\ in\ text\ units\ containing\ accounting\ information}{Total\ number\ of\ words\ in\ the\ report}$$

The independent variables in the model consist of firm, analysts and report characteristics discussed in the previous section. Definition and measurement of each of the variables are discussed below:

The first hypothesis tests the association between analysts' use of accounting information in their reports and the financial performance of the company covered. The level of percentage change in EPS is used to proxy for financial performance. In prior studies, several proxies have been used to measure financial performance or profitability including *return on equity (ROE)* and *return on assets (ROA)* (Li, 2008; Abhayawansa, 2010; Rogers *et al.*, 2011). However, these are rarely cited in analysts' reports. Evidence from prior research (see for instance Previts *et al.*, 1994) and my own reading and coding of analysts' reports reveal that the most cited measure of profitability is *EPS*. Hence, percentage change in EPS is used as a measure of financial performance in this study. This is measured as:

$$P\_EPS = \frac{(EPS_{2010} - EPS_{2009})}{EPS_{2009}}\%$$

Hypothesis H_2 is concerned with the association between analysts' use of accounting information and volatility of earnings, which is used to proxy for operational risk. There are several proxies for risk in extant literature such as volatility of returns (Orens and Lybaert, 2010) and beta (Abhayawansa, 2010). However, in order to capture risk related to company's operations, earnings' volatility is measured, following Li (2008) and Dichev and Tang (2009) as the standard deviation of operating earnings (EBIT) over the five years to 2010, scaled by average assets over the same period:

$$Risk = Standard\ Deviation\left(\frac{Earnings\ Before\ Interest\ and\ Tax\ (EBIT)}{Average\ Asset_{2006-2010}}\right)$$

The third hypotheses is concerned with the effect of firm size on analysts' use of accounting information. Several proxies for firm size have been used in prior research including *revenue*, *number of employees* (Patelli and Pedri, 2013) or total assets (Cho *et al.*, 2010). These are often used interchangeably with none

confirmed superior. Thus for this study, a widely used measure of firm size is used. Following prior studies (such as Asquith *et al*. 2005, Barth *et al*. 2008, Li 2008 and Kothari *et al*. 2009), firm size is measured as:

$$Size = \ln(Market\ value\ of\ equity_{2010})$$

The fourth hypothesis considers growth opportunity and its association with analysts' use of accounting information. Two variables are used to test these hypotheses. Charitou *et al*. (2011) posit that firms that have increasing earnings and/or revenues faster than their industry or overall market are considered growth oriented. Thus the first variable used to measure growth opportunities is the compound average growth rate in sales over the previous five year period (2006-2010), following Lehavy *et al*. (2011). This is computed as follows:

$$SGrowth = \left(\frac{Sales_{2010}}{Sales_{2006}}\right)^{1/5} - 1$$

Several studies have also used market – to – book ratio as a proxy for growth opportunities (see for instance, Li, 2008; Ghosh and Moon, 2010). Ball and Shivakumar (2008) argue that market to book ratio represents 'unbooked' growth options as well as "unbooked" intellectual property of mature non-growth companies, while Charitou *et al*. (2011, p.164) suggest that this variable is only indirectly linked with growth options and have no "theoretical underpinning". However, given its widespread use within accounting literature, for robustness, it is included as a second variable to proxy for growth opportunities. In this study, it is measured as:

$$MTB = \frac{Market\ value\ of\ equity_{2010}}{Book\ value\ of\ equity_{2010}}$$

The impact of financial leverage on use of accounting information by analysts in their reports is tested in hypotheses H_5. Following Kothari *et al*. (2009), leverage is defined as long-term debt to total assets of the firm.

$$LTD = \frac{Long\ term\ debt_{2010}}{Total\ assets_{2010}}$$

Hypotheses H_6, test the extent to which reliance on intangible assets impact on analysts' use of accounting information. Following prior studies (e.g.

Abdolmohammadi *et al*. 2006), *industry classification* is used to measure the extent of reliance on intangible assets. For this study, industry classification is based on Francis and Schipper (1999) classification. The high-technology sample includes industries in which value is mostly generated from intangibles such as patents, copyrights and intellectual property, while the low-technology sample include traditional industries which generate value principally from manufacturing process and are thus more reliant on tangible assets.

Table 1: Industries Included in High- and Low-Technology Samples

High-Technology Industries	
283	Drugs
357	Computer and Office Equipment
360	Electrical Machinery and Equipment, Excluding Computers
361	Electrical Transmissions and Distribution Equipment
362	Electrical Industrial Apparatus
363	Household Appliances
364	Electrical Lighting and Wiring Equipment
365	Household Audio, Video Equipment, Audio Receiving
366	Communication Equipment
367	Electronic Components, Semiconductors
368	Computer Hardware (Including Mini, Micro, Mainframes, Terminals, Discs, Tape, Drives, Scanners, Graphics Systems, Peripherals, and Equipment)
481	Telephone Communications
737	Computer Programming, Software, Data Processing
873	Research, Development, Testing Services

Low Technology Industries	
20	Agricultural Products-Livestock
160	Heavy Construction, Excluding Building
170	Construction-Special Trade
202	Dairy Products
220	Textile Mill Products
240	Lumber and Wood Products, Excluding Furniture
245	Wood Buildings, Mobile Homes
260	Paper and Allied Products
300	Rubber and Miscellaneous Plastics Products
307	Miscellaneous Plastics Products
324	Cement Hydraulic
331	Blast Furnaces and Steel Works
356	General Industrial Machinery and Equipment
371	Motor Vehicles and Motor Vehicle Equipment
399	Miscellaneous Manufacturing Industries
401	Railroads
421	Trucking, Courier Services, Excluding Air
440	Water Transportation
451	Scheduled Air Transportation, Air Courier
541	Grocery Stores

The table is extracted from Francis and Schipper (1999) who examined the difference in value relevance of accounting information across high and low tech firms. The table reports the three-digit SIC codes and the names of the industries included in the high- and low technology samples. According to Francis and Schipper (1999), industries are selected for inclusion in the high- (low-) technology sample based on whether firms in the industry are likely (not likely) to have significant unrecorded intangible assets.

To test hypotheses H_7, reports are grouped into two sub-samples. One sub-sample contains reports issued by IB-analysts and another sub-sample containing reports issued by analysts employed by IND-analysts. Similar distinction has been undertaken in previous studies. For multivariate analysis, a dummy variable *IB* is used to differentiate the two sub-samples and takes the value of 1 for reports authored by IB-analysts and 0 for reports authored by IND-analysts.

Following De Franco and Zhou (2009), analysts' reports are split into two sub-samples. The CFA sub-sample contains reports issued by analysts, of which at least one is CFA qualified (i.e. has the CFA designation after his/her name) and the non-CFA sub-sample contains reports in which none of the authors have a CFA designation after their name. For multivariate analysis, a dummy variable (*CFA*) is used to differentiate the two sub-samples and takes the value of 1 for reports from the CFA sub-sample and 0 for the non-CFA sub-sample.

For univariate analysis, reports are split into three categories based on their recommendation types. This study categorises stock recommendations as *positive*, *neutral* or *negative*. Hence, reports are split into three categories based on whether the accompanying recommendation is positive, neutral or negative. Additionally, for multivariate analysis, a dummy variable (*POS*) is used to distinguish between positive and non-positive recommendation. In this case, reports in the negative category are included in the neutral category. This grouping is in line with prior studies (see for instance, Bradshaw 2002, Abhayawansa and Guthrie 2012).

4.6 Chapter summary

This chapter presents the development of hypotheses used to address the first research objective of this study, i.e. examination of the factors that influence analysts' use of accounting information. First, it presents and briefly discusses hypotheses related to firm characteristics, i.e. percentage change in EPS, volatility of earnings, firm size, sales growth, market-to-book ratios, leverage and industry classification. This is followed by hypotheses related to analysts' characteristics and recommendation types.

Additionally, the statistical model used to test the hypotheses was presented and the choice of measures for the independent variables used in the model was discussed.

Chapter 5: Content Analysis Method

5.1 Introduction

This chapter discusses the application of content analysis to collect data for this study. As acknowledged in Weber (1990, p.13), "there is no simple right way to do content analysis. Instead investigators must judge what methods are most appropriate for their substantive problems", hence, this chapter lays out the practical choices deemed most suitable to address the first research objective. Precisely, section 5.2 introduces the content analysis method, the theoretical underpinnings of the method and explains how it is operationalised in this research; Section 5.3 details the sampling choices and justification and the sample composition; Section 5.4 enumerates the coding process, including details of the recording and measuring unit, issues related to establishing reliability and validity of the coding process; Section 5.5 concludes.

5.2 Content analysis

5.2.1 Definition of content analysis

Berelson (1952, p.18 in Holsti, 1969) defined content analysis as a "research technique for the objective, systematic and quantitative description of the manifest content of communication". This position is rather narrow and restricts content analysis to the analysis of manifest content on the basis of quantification. A broader definition that encompasses all variants of content analysis was adopted by authors such as Holsti (1969, p.14) who defines content analysis as "any technique for making inferences by objectively and systematically identifying specified characteristics of messages". Similarly, Weber (1990) defined content analysis as a method which allows a systematic and objective classification of data in to pre-determined information categories, from which inferences can be made.

Two important features can be deduced from these definitions: (i) objectivity of the process; and (ii) systematic approach to the analysis. Essentially, the research must be carried out on the basis of explicit rules and procedures which are applied consistently (Holsti, 1969). This enables classification of content analysis as a scientific tool of enquiry (Krippendorff, 2004). Moreover, Merkl-Davies *et al.*

(2014) argues that the objectivist epistemology at the basis of the positivist research paradigm to which content analysis is classified, necessitates that the analysis is rigorous, based on explicit rules and procedures, which minimises the subjective analysis of the authors. These attributes distinguishes content analysis from other forms of textual analysis such as *discourse analysis*, in which analysis assumes an interpretative perspective and is based on the use of language to construct reality rather than describe an existing reality. Moreover, it is highly subjective and dependent on the researcher's beliefs and values.

5.2.2 Approach to content analysis

In practice, content analysis involves categorising text into pre-determined categories for analysis and inferences (Guthrie *et al.,* 2004, Steenkamp and Northcott, 2007, Abhayawansa 2010). However, the research design may take different forms and the suitability of each design is contingent on the research objectives. The approaches are distinguished and the choice for this study is highlighted in view of the first research objective of this study.

Based on the content of the text being analysed, Jones and Shoemaker (1994) distinguishes between *thematic* and *syntactic* content analysis. The former being directed at examination of the themes present within a text, while the latter is concerned with linguistic aspects of the text, such as readability. Both the thematic and syntactic aspects of analysts' reports are investigated in this study. However, the first research objective, which this chapter addresses, concerns the thematic analysis of the content of analysts' reports. Syntactic analysis is used to address the second research objective of this thesis and the research design is presented and discussed in Chapter 9.

Based on the measurement approach, content analysis could be *volumetric* or based on an *index* (Vourvachis, 2007). The index approach is concerned with whether a particular word or theme is present in the document being analysed. This is most useful for descriptive studies where the volume of disclosure is not the variable of interest but rather the presence or absence of information items. Importance of each type of information is based on the number of documents that contains it (see for instance, Flostrand, 2006; Orens and Lybaert, 2007 and Garcia-Meca and Martinez, 2007). One of the advantages of this approach, as highlighted

in Vourvachis (2007), is that it is not influenced by the authors' writing style. However, the drawback is that it does not account for the differences in the frequency of occurrence of each information type. On the other hand, volumetric studies are concerned with the quantity of disclosure. Hence, volume of disclosure is derived from a frequency count of the occurrence of each information unit. This approach is based on the premise that volume of disclosure is an indication of importance and that all units have equal weight (GAO, 1989). This is particularly suited for determinant studies in which variation in the volume of disclosure is of interest. Hence, given the interest in measuring the extent of use of accounting information and how this varies across different reports, a volumetric approach is most suitable and thus adopted.

Based on meaning of the words presented in a text, content analysis could be *manifest* or *latent* (form-oriented or meaning-oriented in Smith and Taffler, 2000). Manifest content analysis is based on the "surface meaning of the text" (Holsti, 1696 p.12). The concern is with the literal and denotative meaning as opposed to latent content analysis which is concerned with the connotative meaning of text. Latent content analysis goes beyond the surface and delves deeper into the meaning of words by "reading between the lines" (Abhayawansa 2010, p.113). Manifest content analysis is used in this research to ensure comparability with prior studies of analysts' reports which mostly rely on this approach. Moreover, the choice of manifest content analysis reduces the subjectivity that is intrinsic of latent content analysis and enhances the reliability of the coding.

Based on extent of reliance on computer for coding, content analysis could be *automatic* or *manual*. Automatic content analysis involves reliance on content analysis software to classify text into different categories based on pre-specified wordlists (for instance, Fogarty and Rogers 2005; Kothari *et al*. 2009; Mokoaleli-Mokotelli *et al*. 2009). The researcher's input in the coding process is minimised. On the other hand, manual content analysis is based on human coding. Text portions are coded into different categories by the researcher using pre-specified coding rules (for instance, Abhayawansa and Guthrie 2012). For this study, both automatic and manual content analysis are used. The first research objective is

addressed using manual content analysis[16] and this is discussed in the rest of this chapter while automatic content analysis is used to address the second research objective and the research design is explained in chapter 9 of this thesis.

Following coding, the approach to analysis could be *descriptive* or *determinant*. A descriptive study of the content of a document is one in which the intention is to provide a general overview of its content and describe practice whereas a determinant study explores the associations between the content of the message and other variables in order to explain observed practices (Beattie, 2014) or make inferences about determinants or consequences of the content of a document. In this study, a determinant form of analysis is adopted, following coding. This enables the association between the content score (which measures the extent of use of accounting information in analysts' reports) and firm, analysts and report characteristics to be explored, such that inferences can be made about factors which explain variability in the use of accounting information in the reports.

16 In the first instance, a semi-automatic content analysis was contemplated as this enabled coding of more analysts' reports than the manual content analysis, given its time-consuming and laborious nature. This was to involve creation of a list of accounting keywords and subsequent use of computer program to automatically search analysts' reports and classify the keywords into different categories. The automatic classification was to be accompanied by a manual check of the context of use of the keywords and setting of the boundary for the text portions. However, during the test coding, it was evident that analysts often used esoteric terms and abbreviations to refer to financial statement elements, thus the list of possible keywords is inexhaustible. Moreover the use of pronouns to refer to previously mentioned accounting information was also pervasive and not likely to be captured with the proposed approach. Hence, to ensure validity and limit both type I and type II errors, manual content analysis was used.

5.3 Data

This section discusses the sample of reports used in this study. The issues addressed are discussed in separate subsections as follows: section 5.3.1 introduces the sampling unit and justifies this; section 5.3.2 provides details about additional sampling choices and section 5.3.3 explains data sources, the data selection process and presents the sample composition.

5.3.1 Sampling unit: analysts' reports

The choice of a sampling unit is one of the first decisions to be made when undertaking content analysis (Vourvachis 2007). Sampling units are defined as "units that are distinguished for selective inclusion in an analysis" (Krippendorff 2004, p.98). Though, there are various means through which analysts disseminate their views about the companies they follow, such as emails and telephone communications with clients, interviews, articles and commentaries, nevertheless, these sources of analysts' communication are either generally not publicly accessible as the reports (Abhayawansa 2010) or their existence is not known. Hence, equity research reports are used as the sampling units for this study.

Sell-side analysts' reports are not only publicly available[17] but possess certain features which make them suitable for this study. First, the reports provide a useful document to understand and make inferences about analysts' work in general. Breton and Taffler (2001, p.92) argue that the reports are the "only extensive trace of analysts' work". However, it is also well-acknowledged that the reports do not offer details of analysts' decision processes (Govindarajan 1980) as not all information used by financial analyst may be contained therein (Schipper, 1991). Nevertheless, the CFA Institute requires that "members and candidates should communicate in a recommendation the ***factors that were instrumental*** in making the investment recommendation" (CFA Statement of Practice Handbook, 2010; p.137 emphasis added). Hence, it is expected that the reports

[17] Generally the reports are obtained from financial databases such as Thomson Financial and access to these databases are based on payment of subscription fees. Hence, while the reports are publicly available, access is at a cost.

contain information used by analysts in arriving at their recommendations. Second, the suitability of analysts' reports is also marked by the importance of the reports for investors and their influence on market variables. The reports are relied on by several groups of capital market participants whose decision processes are dependent on the information contained therein (Fogarty and Rogers, 2005; Campbell and Slack, 2008). Empirical evidence suggests that the provision of value relevant information in the reports is also marked by significant market reaction to their release (Asquith *et al*. 2005). Consequently, the reports contain information that influences the wealth distribution process.

5.3.2 Sampling choices: analyst-firms, companies and reports.

Unlike companies' annual reports which are issued once a year by the company, there are several analysts' reports on any given company each year. These are issued by several analysts covering the company and are issued at different times during the year, following different events. Thus, the next step in the sampling process is to identify which analysts' reports to include in the sample for this study. There are three key areas in which sampling choices are to be made: (i) the analyst-firms, (ii) the analysts' reports and (iii) the companies. These choices are discussed in the following sub-section.

5.3.2.1 Analyst-firms

Analysts' reports are authored by analyst(s) working for an investment bank/brokerage house or independent research firm. These analyst-firms have differing features such as size, international presence, nature of operations etc. However, for this research, reports are randomly selected across analyst-firms in order to ensure the results are reflective of a broad range of analyst-firms.

As previously mentioned in preceding chapters, a focal point of this research is the distinction between reports issued by IB-analysts and IND-analysts. Hence, two analysts' reports are selected for each company, one being issued by an IB-analyst and another being issued by an IND-analyst. The features of IB-firms and IND-firms have been previously discussed in Chapter 2 of this thesis. Classifications of analyst-firms into these categories were made following previous studies. Jacob

et al. (2008), Barber *et al*. (2007), Cowen *et al*. (2006) and Gu and Xue (2008) categorises analyst-firms into investment banking and non-investment banking or independent firms using Nelson's directory of investment research. The directory provides a profile for over 14,000 analyst-firms and classifies them as "Investment Bank/Broker", "Investment Manager" and "Independent Research Firm". The 2008 version[18] of the directory was obtained and list of independent firms were distinguished from Investment banks and brokers. To deal with possible switches of analyst-firms between categories and for firms not listed in the directory, firms' website and the appendices of their reports which contains regulatory disclosures were further examined (as in previous studies such as Cowen *et al*. *(*2006) and Barber *et al*. (2007)). Firms are classified as IB-firms if they offer investment banking business and/or brokerage services while IND-firms only engage in research and do not offer these services. In searching websites and regulatory disclosure notes, firms classified as independent research firms make comments such as these on their websites and/or the appendix of their reports:

> Argus Research is not a registered broker dealer and does not have investment banking operations.
>
> Indigo Equity Research Limited has no investment banking or share trading operations or activities.
>
> Pivotal Research Group LLC is an independent equity research company and is neither a broker dealer nor offers investment banking services. Pivotal Research Group LLC is not a market maker for any securities, does not hold any securities positions, and does not seek compensation for investment banking services.
>
> Wall Street Strategies is not a broker/dealer, and the firm does not underwrite securities, manage assets or provide investment banking activities.

The list of analyst-firms used in this study is presented in Appendix A.

[18] The 2008 version was the latest version obtainable. Further internet search reveal that the directory may no longer be in production.

5.3.2.2 Companies

The firms analysed in this research are drawn from the S&P 500 index. The S&P 500 index is an equity market index which contains 500 of the leading stocks in the US by market capitalisation. The constituent list as at 31 December 2011 and their respective four digit SIC code were obtained from the Computstat database. This group of firms were chosen for three main reasons. First, their larger size results in greater availability of analysts' reports on *Investext* as larger firms generally attract more analyst coverage (Lang and Lundholm, 1996; Lehavy *et al*. 2011). Second, the constituents represent about 75% of the market capitalisation of the US market. Hence, they are most influential and representative components of the market. Third the US setting is justified given its strong international influence in the area of capital markets development and financial analysis. For instance, the current approach to financial analysis is significantly influenced by the seminal writings of the American, Benjamin Graham, who is often referred to as the as the father of modern financial analysis (Muller, 1994). In his famous book, *Security Analysis* (co-authored with David Dodd), Graham proposed a rigorous framework for analysing a firm's future earnings and cash flow potential.

The US influence on the practice of financial analysis is also supported by the Chartered Financial Analysts (CFA) Institute, which is the professional body for financial analysts across the globe. Birthed by the US Financial Analyst Federation and influenced by the writings of Graham, who also pioneered calls for a professional rating for security analysts, the CFA Institute sets standards for the practice of financial analysis and certifies analysts upon completion of their qualification. Furthermore, several international brokerage houses originate and are headquartered in the US, the US capital markets is well developed with several exchanges, it has a strong presence of sell-side analysts, who have gained popularity as an influential group of capital market participants. Finally, the issue of independence and objectivity of sell-side analysts has been very much discussed and regulated in the US, particularly after the burst of the dotcom bubble. Moreover, the choice of S&P 500 constituents enable comparison with previous studies (e.g. Rogers and Grant 1997; Flostrand, 2006 and Flostrand and Strom 2006).

Additional selection criteria are that the sample is limited to non-financial companies. There are several reasons for excluding financial companies. First, financial companies are highly regulated, which influence their overall financial reporting and disclosure practices. Thus, analysts' use of accounting information for these companies might be partly explainable by their different regulatory environment. Second, financial companies have highly geared capital structures. Hence, the effect of varying level of leverage on the use of accounting information may not reflect the arguments advanced in prior studies (as discussed in Chapter 4) regarding the effect of gearing on the value relevance of accounting information. Third, given the nature of their operations, the classification of financial statement items differs in some respect to non-financial companies. For instance, while loans are generally classifiable as liabilities for a non-financial company, loans are classified as assets for financial companies such as banks. Such differences may limit the applicability of the coding rules to classification of accounting information for financial companies. Finally, exclusion of financial companies facilitates comparison with prior studies of the content of analysts' reports such as Rogers and Grant (1997), Breton and Taffler (2001) and Orens and Lybaert (2007).

5.3.2.3 Analyst reports

For each company in the sample, multiple reports are issued during any given year by several analysts. The general observation from prior literature on the content of analysts' reports is that reports type has not been clearly distinguished. Nevertheless, Garcia-Meca and Martinez (2007) showed that type of reports is an influential variable when studying the content of analysts' reports. Hence, prior to selecting the type of research reports to be used in this analysis, previous empirical studies were examined as well as the *Investext* database which was used to source the reports. The aim was to review the types of reports that are generally available and identify the most suitable for this study.

Only a few studies comment of the type of analysts' reports used in their research (Flostrand, 2006; Garcia-Meca and Martinez, 2007; Abhayawansa and Guthrie, 2012 and Twedt and Rees, 2012). Some of these authors not only differentiate between the types of report but also examined the effect of types of reports on their analysis. For instance, Garcia-Meca and Martinez (2007) observed how

intellectual capital related information varied across results reports and company reports, with results reports defined as reports issued following quarterly, interim or annual results and company reports defined as reports issued following a change in a firm's strategy or formation of new alliance. Their findings suggest that the information disclosed in the reports varies across types of reports, with intellectual capital information being significantly higher in company reports compared to results reports.

In general, three types of reports are mostly referred to in extant literature. A *fundamental* or *company* report, *result* report and *initiating coverage* report. These reports are distinguished based on the event that motivates the issuance (results reports are issued following results announcements and initiating coverage reports are issued following a decision to follow a company) or the depth of analysis (e.g. fundamental reports are more comprehensive).

Analysts issue equity research reports for a variety of purposes and events. Through a review of the *Investext* database the following types of reports are generally found to be available.

By Author

Equity analysts: Equity research reports are issued by sell-side analysts[19]. They generally contain investment recommendation, forecast of earnings and other financial metric, a price target and analysis/justification of the recommendation.

Debt analysts: These reports are written for debt valuation purposes usually by credit analysts and generally contain debt recommendations and forecast of debt returns (Johnston *et al*. 2009). While being relatively similar to equity research reports, the object of the analysis is debt rather than equity.

Others: These reports generally provide analysis of a company and the business environment in which they operate and/or their risk profile. Usually, these reports

[19] Buy side analysts also issue equity reports, however these are mainly for their fund management firm and not generally available to the public through financial databases such as *Investext*.

do not contain any investment recommendation or forecasts. They are generally authored by risk analysts or strategists rather than equity or debt researchers.

By Event

The process of analysing a company's worth and projection of future cash flows is not static. During the course of any given year, there are several corporate or external events that influence analysts' views and previously held beliefs about the companies followed. Hence, the process of analysis is ongoing till ceasing of coverage. From a review of the *Investext* database, the occurrence of several events is generally accompanied by the issuance of analysts' reports. These include:

Initiating coverage/ceasing coverage: These are reports issued following analysts' decision to begin or terminate coverage of a company. The initiating coverage reports are often very detailed and more comprehensive as they are "concerned with forming a new knowledge base as opposed to other types of analysts' reports, which are considered to be merely updates on an existing base of knowledge" (Abhayawansa 2011, p.452).

Result announcement: These are reports issued following the announcement of a company's quarterly, interim or annual financial results and ensuing conference calls. The results reports are recurrent and appear to be standard practice across firms and analysts. Another type of report that could be grouped under this category is the pre-result reports. These are reports concerned with the firm's results' announcement. However, they are issued prior to the event and are generally aimed at forecasting/anticipating the results. These are relatively few compared to post-result reports.

Company visits: These are reports issued by analysts following a visit to the company being followed. They often disclose information obtained from management during such meetings and any prior forecast or recommendation are updated based on any new information. Given that these reports are issued following company visits, the frequency and timing of which differ across companies, they are rather sparse.

Morning notes: These are very short update reports issued following any new information.

Ad hoc company events: These reports are issued following firm specific events such as any new equity offering, mergers and acquisitions, new product, change of strategy etc. The list of event is inexhaustible and differs across companies. Hence, it is difficult to predict the timing and availability of these reports across companies.

By Object of Analysis:

Multi-company reports: These reports analyse different companies simultaneously. They are often focused on comparison of a company's performance and outlook relative to its peers. This can also take the form of an industry analysis in which the companies within the sector are analysed.

Single Company Reports: These reports are issued for a company at a time. Most reports fall within this category.

For this research, only reports by equity researchers are included in the sample, consistent with this study's focus on sell-side analysts. Reports by both investment banking and independent research analysts are used. Furthermore, to ensure comparability across companies in the sample, reports which are issued following certain company specific, ad-hoc events, company visits and morning notes are not considered as these events are not predictable nor are they simultaneously issued across companies. Hence, the results and initiating coverage reports are plausible. However, on further search of *Investext* database across several companies from the pilot sample, there was a dearth of initiating coverage reports in the year of interest. Moreover, it was rare to find initiating coverage reports issued by both a sell-side and independent research analyst simultaneously. Consequently, the sample consists only of *results* reports. Finally, only *results* reports issued about a single company are included in the sample and multi-company or industry reports are excluded. This exclusion is necessary as the objective of this research is to examine the factors that explain the variation of the content of analysts' reports. Factors considered include firm-specific, such as profitability, size and growth characteristics. Hence, this analysis cannot be

achieved through multi-company reports as the companies discussed in such reports will generally have different attributes.

5.3.3 Source of data and sample selection process

5.3.3.1 Source of analysts' reports

The analysts' reports used for this study were obtained from Thomson's Investext database. *Investext* describes itself as the world's most comprehensive database of analyst research reports with over 12 million research reports from over 1,600 investment banks and independent research firms with more than 30 of these being exclusive to *Investext* such as Credit Suisse, Morgan Stanley, and HSBC Global Research. *Investext* has been widely used as a source of analysts' report by several content analytic studies (such as Bradshaw 2002; Asquith *et al.,* 2005). Despite its wide coverage, certain brokerage houses and investment banks do not make their reports available through *Investext* such as Goldman Sachs. Nevertheless, new contributors are constantly being added from all over the world and currently include reports from 9 of the 10 top rated institutional investors' All-America team 2011 and 17 of the top 20 firms in the Extel Pan European Survey 2011 (Thomson Reuters, 2012).

5.3.3.2 Report selection process and sample composition

Following from the sampling choices discussed in section 5.3.2, reports included in the final sample must meet the following criteria: (i) be based on equity research; (ii) cover a single company; (iii) be issued in 2011[20], immediately following the announcement of the 2010 or 2011 annual results; (iv) be issued before the end of the first quarter of the next fiscal year, and (v) be issued by analysts employed by an IND-firm or an IB-firm[21].

[20] Download of reports begin in 2012, thus, 2011 was the most recent year to include in the sample. Moreover, a single year as selected due to the labour intensive nature of manual content analysis. Also, to address the research objective, cross-sectional variation in the use of accounting information was of interest rather than variation over time.

[21] For each company, two reports were downloaded from *Investext*, one by analysts employed by an IB-firm and another by analysts employed by an IND-firm. Hence, companies with reports issued only by an IB-firm are not included in the sample. This criteria is necessary to address the research question, as it provides a matched sample for investigating the variation in content of analysts' reports across both types of analysts covering the same companies,

The reports selection process included several steps:

- First step involved obtaining the list of S&P constituents as at December 2011 and their SIC codes. This was obtained from Computstat database.

- The second step involved eliminating financial companies from the list. These are firms with SIC codes 6000-6999. Rational for excluding financial companies has been previously discussed.

- Because results reports are the object of this study, for each of the companies, the annual results announcement and first quarter results announcement date were obtained from the corporate event section of ThomsonOne database. This facilitates location of annual results reports as they are often released immediately after earnings' press release at the end of the quarter. Moreover, it is comparable with other studies such as Rogers and Grant (1997) and Orens and Lybaert (2007).

- For each company, *Investext* database was searched for a list of analysts' reports published between the above-mentioned dates.

- For each company, two reports were randomly selected – one by IB-analysts and the other by IND-analysts. Companies with reports issued by only one type of analysts (e.g. IB-analysts) were excluded from the sample.

Using the above criteria and following the above steps resulted in a total sample of 288 analysts' reports for 144 firms. The sample size is larger than those used in most studies on the content of analysts' reports such as Fogarty and Rogers (2005), which used 187 reports and Abdolmohammadi *et al*. (2006) and Abhayawansa (2010), which used less than 70 reports. Table 2 presents a summary of the sample selection procedure.

---

which is the main focus of the second research objective and also of interest for the first research objective of this thesis.

Table 2: Sample selection

S&P constituents	500
Less financial companies	85
Less companies with missing independent analysts reports or results reports for 2011	271
Final sample of companies	144
Final sample of reports (2 reports per company)	**288**

5.4 Content analysis process

This section lays out the actual steps involved in content analysing the reports. These steps are based on the approaches to developing and applying a coding scheme suggested by Weber (1990) and Boyatzis (1998). Table 3 provides a summary of the main issues addressed and the sections in which these are discussed extensively.

Table 3: Content analysis process

Content analysis process	Section
Defining recording and measuring unit	5.4.1
Defining categories and coding rules	5.4.2
Assessing reliability and validity	5.4.3

5.4.1 Defining recording and measuring units

In content analysis literature, several types of units are often referred to by researchers, including *unit of analysis*, *context units*, *coding units etc*. For the purpose of this research and following the suggestions of Weber (1990) and Boyatzis (1998), the main units of interest discussed and analysed are the *recording* and *measuring* units. Vourvachis (2007; p.13) highlighted the variety of terminologies that has been used in existing literature to describe units in content analysis. He noted that "... Milne and Adler's (1999) *coding units* are what Krippendorff (2004) describes as *context units* and Neuendorf (2002) as *analysis units*, whilst Milne and Adler's *measurement units* are Krippendorff (2004) *recording/coding units* and Neuendorf (2002) *data collection units*". For the purpose of this research, the terms *recording units* and *measurement units* are used. Holsti (1969; p.116) defines recording units as the "specific segment of content that is characterized by placing it in a given category". This is similar to GAO (1989; p.10) definition of a recording unit as the "specific segment of the context unit in the written material that is placed in a category". As the central feature of content analysis involve the splitting of text into several units for classification and further analysis, *recording units* represent portions of text which are assigned to pre-specified categories. In other words, the *recording unit* represents the unit of the text to be coded (Weber, 1990). Abhayawansa (2011;

p.454) defines the *measurement unit* as the "basis for quantifying the amount of disclosure". In essence, the *recording unit* forms the basis for coding, while the *measurement unit* forms the basis for measuring the amount of disclosure of an information type (Milne and Adler, 1999; p.243). Most often these units are the same but they could be different.

The importance of selecting an appropriate recording and measuring unit has been emphasized in existing literature. For instance, Unerman (2000; p.674) noted that "different measurement techniques might lead to different impression of the relative importance of each disclosure category". Hence, the choice of the most appropriate unit requires thoughtful consideration, given its importance in shaping the results of the analysis. A variety of recording and measuring units have been used in previous accounting research. These are discussed in the following paragraphs[22], their features are highlighted and the most appropriate units are selected for this study.

5.4.1.1 Words/Terms

Words are the smallest unit of text that can be analysed. The analysis will often involve the classification of keywords into several categories and measurement is based on the number of keywords in each category. A variant of the use of words is represented by *terms,* which is a combination of words that have a specific meaning (Abhayawansa, 2010) *e.g. employee training*. *Words* or *terms* are mostly used by authors who make use of computer assisted content analysis programs in their analysis. These programs, such as Diction, Linguistic inquiry and Word Count (LIWC) and General inquiry, have specific dictionaries of keywords and the researcher can also create a custom dictionary of words.

For an example, Barker and Imam (2008) used keywords to examine the content of analysts' reports. They sought to understand whether earnings quality is perceived by analysts from an accounting or non-accounting perspective. Each

[22] The review of the units used in previous analysis is limited to accounting literature with emphasis on studies which examined analysts' report narratives. The rational for this is that recording units varies across subject areas. For instance, Holsti (1969) identified 'Characters' as a recording unit for studies of fiction, drama, movies. This is not particularly relevant for this study and its discussion will only deviate from the scope of this research.

keyword was categorised into accounting or non-accounting based on its context. Similarly, Breton and Taffler (2001) examined the relative importance of 5 different themes (Growth, Management and Strategy, Profitability, Financial Position and Market Conditions) in explaining types of stock recommendation. Each theme was represented by several keywords used in the context of the thematic categories. The content analysis approach involved generating thematic scores for each report analysed, which was computed as the frequency of occurrence of the keywords in each theme, divided by the total number of words in the text.

The main advantage of using *words* or *terms* as a recording unit is the reliability of the coding process (Krippendorff, 2004). Once the dictionary of keywords is identified by the researcher, the coding requires reading through the text to find an occurrence of the words, which is generally straightforward for form-oriented content analysis. However, the use of words as a recording unit presents several limitations. First, it has been argued that due to the different meanings of a single *word*, the use of *words* as a recording unit poses a threat to the validity of the research. It is important that the meaning of each word be understood based on the context of its use within a sentence (Milne and Adler, 1999)[23]. Second, within the context of narratives, the use of *words* ignores the differences in formatting present in narratives such as *boldness, font, size*, which are all features useful in conveying importance of the information disclosed (Beattie and Thomson, 2007). Third, with regards to measuring the extent of disclosures, studies that use word count generally compute the extent of disclosure as the frequency of words in a particular category divided by the total word count in the report. However, this practice does not fully reflect the space each information type occupies in the document. This issue is illustrated by the following extract:

> "PTR's **growth** forecast for Google begins with consideration of historical **growth** rates. Four historical **growth** measures are examined for up to the last ten years. Each **growth** measure is derived on a per share basis so that dilution from acquisitions or financing is reflected. **"Balance Sheet" growth** (**growth** in **assets** or **equity per share**) is

[23] A keyword in context analysis (KWIC) may overcome this limitation as it is based on analysis of the context of use of the keywords.

> typically the best measure of a company's basic -- or intrinsic -- **growth** rate."
>
> *(Source: Price Target Report on Google Inc, February 17th 2008)*

The total word count in the above extract is 69 and the number of *growth* words is 7. In a situation in which words are used as the unit of recording and measurement, the importance of the growth category will be computed as 7/69 (0.10 or 10%). This will indicate that growth only occupies 10% of the authors' statement. However, the entire paragraph is about *growth*. The other 90% words are used to describe the concept of *growth* as intended by the analyst. Hence, the use of words could paint an incomplete picture of the importance of each information category. The issue could be further exacerbated in a situation in which the reporting analyst uses pronouns in replacement of keywords. As pronouns could be indicative of any subject matter, they cannot usually be identified as a keyword for a particular category without the context, resulting in missing out valid information from the analysis. Fourth, the use of words both as recording and measuring units has also been criticised for its inability to capture and measure non-narratives parts of a document (Unerman, 2000; Abhayawansa 2011), although this is not relevant for this research.

5.4.1.2 Themes or Text units

Holsti (1969; p.116) defines a theme as a "single assertion about some subject". This has been identified as the most used recording units in analysis of accounting narratives (Jones and Shoemaker, 1994). Although themes and text units are often taken to mean the same unit (as in Abhayawansa, 2010), themes are not a grammatical unit, rather they represent a "conceptual entity" (Carney, 1972, p.159). 'Text units' are defined in Beattie *et al*. (2004, p.216) as a text portion containing a "single piece of information that is meaningful in its own right". For an example, Abhayawansa (2011) and Abhayawansa and Guthrie (2012) used text units as recording units for content analysis of analysts' reports.

The use of text units to record portion of text into categories provides some advantages over use of *words*. First, it captures the words used to convey a single information, thus the context and space occupied by each information type is recorded. Second, given that each text unit contains portion of text regarding a single piece of information, it avoids the problem of multiple coding of same text

portions to different categories. However, a limitation of text units is that, unlike the other units which are defined grammatically, there is no specific boundary for a text unit as this is defined by the researcher.

5.4.1.3 Sentences

A sentence may also be used as the recording unit in content analytic research. Some researchers argue that this is more useful than words, as the meaning of a word is determined by its context, which is often provided by the sentence (Milne and Adler, 1999; Gray *et al.*, 1995). This has formed the recording unit in several content analytic studies in accounting. As with any recording unit, there are advantages and disadvantages in using this method.

The major advantage lies in the ease of identification given it is a grammatical unit with a clear boundary and this enhances the reliability of the coding process. Moreover, it captures the context in which information is used and thus limits misclassification errors. However, the use of sentences is problematic for several reasons. First, Abhayawansa (2011) argued that the use of sentences as recording and measuring units is influenced by writing styles. He observed that same information may be disclosed through different number of sentences and may result in attributing higher frequency of disclosures to documents containing the same information but greater number of sentences. In other words, "using sentences as the unit of measurement seems to ignore the possibility that differences in use of grammar might result in two different writers conveying the same message by using a similar number of words and taking up a similar amount of space but using a different number of sentences" (Unerman 2000, p.675). Second, the use of sentences is not capable of capturing non-narrative disclosures or the format of the text (Unerman, 2000). Third, more than one information item could be discussed in any single sentence thus, the mutual exclusiveness of recording units becomes unachievable (Abeysekera, 2006). Hence, the categories contain non-mutually exclusive units resulting in unit counts beyond the 100%, which could introduce bias in the statistical analysis and interpretations of results.

5.4.1.4 Paragraphs and whole reports

A paragraph or the entire document can also be used as the recording unit. While these units can be easily identified, the limitations discussed with reference to the use of sentences, also apply to these larger units. These units usually contain references to more than one information item, which limits classification into a single category and thus, do not enable mutually exclusive recording units.

5.4.1.5 Choice of recording and measurement unit

All the recording and measuring units discussed above possess certain limitations and benefit. The choice of any unit will therefore be based on considerations of the research objectives being addressed. This section presents the choice of recording and measuring unit and discusses the rationale for their selection.

Table 4 summarise the key features of the units as discussed in the previous section, with the aim of identifying the most important features for this research.

Table 4: Features of recording units

Features	Description	Words/ terms	Text units	Sentence	Paragraph	Importance
Context	Recording unit captures the context in which accounting information is used	No	Yes	Yes	Yes	Very important
Space*	Recording unit captures the space occupied by each information	No	Yes	Yes	Yes	Very important
Mutual exclusivity	Recording unit enables coding to only a single category	Yes	Yes	No	No	Very important
Identifiable	Recording unit is easily identified and distinguished from other units	Yes	No	Yes	Yes	Less important
Formatting features	Recording unit captures the formatting features in the text such as font and size	No	No	No	No	Not important
Non-Narratives	Recording unit is useful for coding visuals and other non-narrative content	No	No	No	No	Not important
* This refers to the space occupies by the accounting keyword as measured by the number of words used in relating the information						

The first six columns identify the features of each type of recording unit and the final column indicates the relevance of each feature for this research project. To obtain a measure of the extent of use of accounting information by analysts (which is the dependent variable in this research), it is necessary that the method of analysis, including the choice of the recording and measuring unit, captures the accounting keyword used, the context of its use and the space occupied within the reports (i.e. the number of words used to convey information about the accounting keyword). The need for this derives from the partial content analysis method adopted for this research. Of all the information contained in analysts' research reports, only accounting information is being investigated. Consequently, failure to include the context and space in the analysis, could result in certain portion of the text (such as function words and pronouns used to convey accounting information) incorrectly classified as not related to accounting information. The process of coding using sentences, text units or paragraphs overcomes this limitation. However, another important feature of relevance for this research project as indicated in the above table is the requirement of mutually exclusive units. In other words, each segment of text should be classifiable into a single category. This is often not possible when large units are used such as sentences or paragraphs. Thus, the recording unit which possess the features most suitable for this research project is 'text units'.

Measurement units form the basis of quantification of information. The use of text units as measurement unit is problematic in this case because it requires the total number of text units in each report[24]. This necessitates the development of coding rules for every information type contained in the reports, which is beyond the scope of this research. Previous studies have used different approaches to overcome this problem. Barker and Imam (2008) used the scaled number of text units per category over total sentences in the reports. The total number of sentences, is easily obtained given that sentences are a grammatical unit separated by the '.'. However, text units are different from sentences and are generally parts of sentences, though they may coincide in some cases. Hence, this approach may result in an inaccurate measurement of the level of disclosure. For

[24] This is necessary as quantification is usually based on total text units containing the information of relevance divided by total text units in the reports.

instance, in a situation in which total number of text units related to accounting information coincides with total number of sentences in the reports, a 100% score will result, which shows a complete concern with accounting information. However, if text units are parts of sentences, such measure inaccurately represents the extent of use of accounting information.

A second approach, adopted in Abhayawansa and Guthrie (2012), is to obtain the total number of text units in a single information category and scale this by the total number of coded text units rather than total text units across each report. For instance, in their study of intellectual capital information, to obtain the proportion of external capital related information, they added the total number of text units containing references to external capital related information and divided this by the total number of text units containing intellectual capital information across all reports. However, such measure only provides an indication of the average use of information related to external capital rather than the use of external capital information in each report. This limits its usefulness for this study because to address the research question, it is necessary to obtain measures that show variability across reports.

Given the limitations of these approaches, this research uses *words* as the measuring units rather than text units. However, unlike studies which are based on keyword counts, the total number of words in accounting text units is used. First, the text units are used as recording units to classify portion of texts into different accounting categories. Second, to measure the extent of use of accounting information and related sub-categories, the total number of words contained in accounting text units for each report are added and scaled by total word count of each report. This provides a measure which shows the extent of use of accounting information in each report and is similar to other studies in the risk reporting literature such as Abraham and Cox (2007).

In summary, the thematic score which measures the extent of use of accounting information involves the use of ‘text units’ as the recording unit and ‘words in the text units’ as the measuring unit and is computed as follows:

$$Extent\ of\ use\ of\ accounting\ information\ per\ report^{25}$$

$$= \frac{No.\ of\ words\ contained\ in\ \ text\ units\ with\ accounting\ info\ (in\ each\ report)}{Total\ no.\ of\ words\ in\ (each)\ report.}$$

The below example illustrates the coding and measuring process as described above:

Extract:

> The Fresh Dairy Direct segment reported sales growth of 5.1% to $2,626 million. Fluid milk volumes were down 3%, but total units for the segment declined 5% due to volume erosion in cultured products, ice cream, creamers and other products. Segment profits were $114.4 million, off significantly from the year-ago figure of $163.7 million but in-line with 3Q10 segment profits of $116.5 million.
>
> *(Source: Northcoast research report on Dean Foods Co., dated February 17th, 2011)*

Table 5: Illustrative example

Text portions	Number of words in text units
The Fresh Dairy Direct segment reported sales growth of 5.1% to $2,626 million.	13
Fluid milk volumes were down 3%,	6
but total units for the segment declined 5%	8
'due to volume erosion in cultured products, ice cream, creamers and other products.	13
Segment profits were $114.4 million,	5
off significantly from the year-ago figure of $163.7 million	9
but in-line with 3Q10 segment profits of $116.5 million.	9
Total Word Count	63
Total words in text units containing accounting information	36
Extent of use of accounting information	57%

Note: Text units containing accounting information are highlighted in grey.

[25] This concords with Weber (1990,p.35) view that "counting words in each category in each document provides a score that measures the intensity of concern with each category in a given document. Counting assumes that higher relative count reflect higher concern with the categories".

5.4.2 Defining categories and test coding

The second step in the coding process is the definition of categories. Holsti (1969, p.95) defines these as the "pigeonholes into which content units are to be classified". The essential feature of content analysis method is the reduction of text (or other communication types) into fewer meaningful categories for analysis and inferences. Hence, the process of category construction is pivotal for the analysis and interpretations. Following the suggestions of Holsti (1969, p.95), the following principles of category construction were adhered to in this research:

Must reflect the purpose of the research: The categories must be based on the variables of interest for the research question.

Exhaustive: All relevant items of content must be capable of being coded into a category. Hence, any reference to accounting information in the analysts' reports must have a suitable category for classification.

Mutually exclusive: This principle requires that the categories should be precise enough to ensure that no unit of text can be placed in more than a single category.

Independent: This requires that assignment of data into a category is not influenced by classification of other data.

Derived from a single classification principle: This principle requires that different levels of analysis be kept separate.

According to the first principle, variables are to be defined both conceptually and operationally to provide a basis for precise and unambiguous coding. As aforementioned, the principal theme being investigated in this research is the use of accounting information in equity research reports. Hence, this section provides a definition of 'accounting information' to be adopted in this study and introduces the related classification scheme which is based on the stated principles.

5.4.2.1 Defining accounting information

In developing a categorisation scheme for accounting information for this research, it is imperative to first define the boundary of accounting information. Prior studies which examined information contained in sell-side analysts' reports have either included accounting information in a broader category of financial information (Abdolmohammadi *et al*. 2006) or financial and operating information (Rogers and Grant, 1997) or have only focused on a sub-set of accounting information (such as earnings' quality in Barker and Imam 2008 or earnings and cash flows in Govindarajan 1980). However, for the purpose of this research, the view is taken that accounting information is but a subset of financial information. For instance, market prices or returns are financial but not accounting information. Moreover, accounting information is not limited to earnings, although this is widely used as a summary measure for accounting information.

The first step in identifying a working definition of accounting information for this research is a review of prior studies. Using the keywords "accounting"/"accounting information", searches were made for studies focusing on accounting information. However, the results reveal that most studies either failed to provide a clear definition of what is meant by accounting information (e.g. Lambert *et al*., 2007) or it is implicitly equated with financial statements information as these terms are used interchangeably (see for instance, McEwen and Hunton, 1999). Often, the definition adopted in such studies is influenced by their research objectives. For instance, Bushman and Smith (2001; p.238) defined (financial) accounting information as the "product of corporate accounting and external reporting systems that measure and publicly disclose audited, quantitative data concerning the financial position and performance of publicly held firms". Further, Orens and Lybaert (2010) examined the use of non-financial information by analysts and defined this as information not presented in the basic financial statements[26], including ratios (with the exception of EPS). This equates financial information with accounting information as contained in the financial statements. An alternative view was provided by Barker and Imam (2008, p.313), who defined non-accounting information as "information drawn from outside the

[26] Financial statements are composed of the Income Statement, Statement of Financial Position, Cash Flow Statement, Statement of Changes in Equity Capital and the accompanying Notes.

financial statements", implying that accounting information is information drawn from the financial statements. This includes uses of accounting information in the forms of ratios, valuation models or forecasts of financial statement information, which are derived directly (ratios and valuation models) or indirectly (forecasts of financial statement information) from the financial statements. Barker (2001, p.45) explained that such uses of financial statement information represents the analytical framework within which investors use accounting information in the context of equity valuation. He further posits that the choice of a valuation model or ratio is an indication of investor's preference for the underlying accounting information. Similarly, the IASB conceptual framework posits that financial information is relevant (i.e. capable of making a difference in decisions made by users) if it has predictive value (i.e. can be used as an input to processes used to predict future outcomes such as valuation models and forecasts). Hence, in assessing the factors that explain the variation in the use of accounting information by analysts, these uses of accounting information needs to be captured.

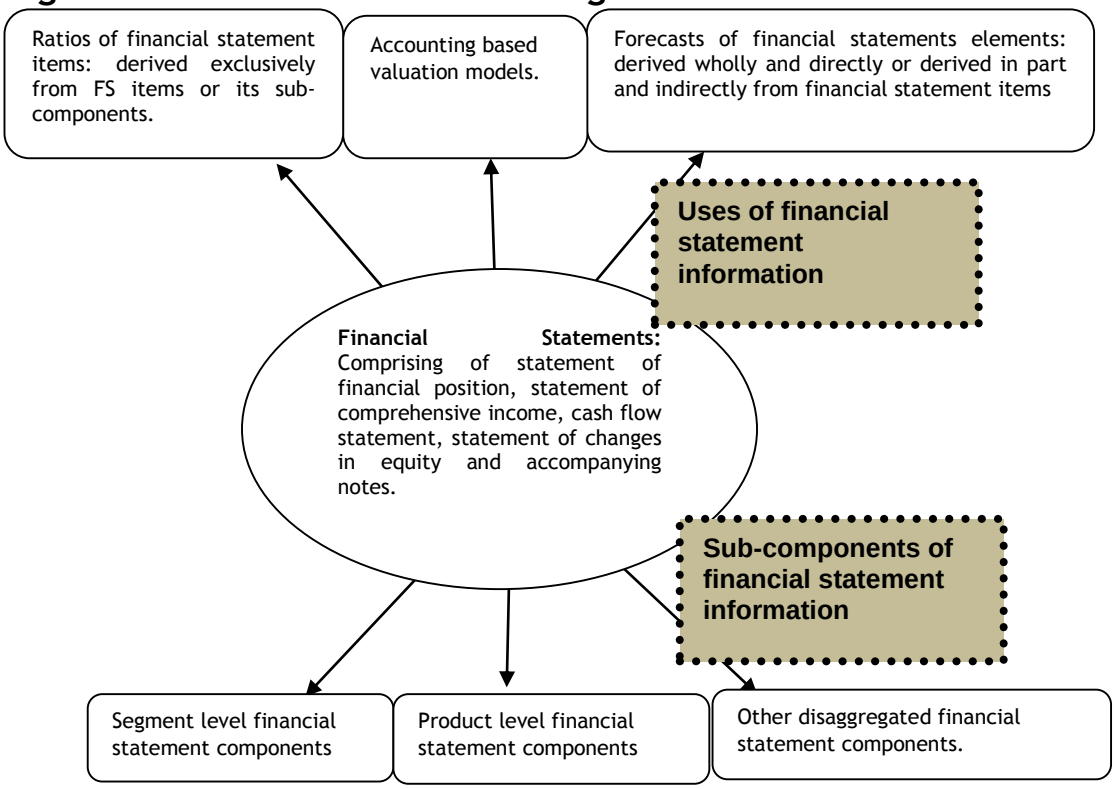
Accounting information has also been defined in publications by accounting professional and regulatory bodies. For instance, the Institute of Chartered Accountants in England and Wales (ICAEW 2009; p.3) considers disclosures of accounting information as financial reporting and distinguishes it from non-financial reporting, which is every other information other than accounting. In a similar vein, the Financial Reporting Council (FRC 2012; p.6) distinguishes between corporate reporting, financial reporting and financial statements. Corporate reporting is referred to as "all the information produced by an entity for users", while financial reporting is considered "information typically found in an annual report, interim or preliminary announcement" and within this subset of reporting, is the financial statements. The Jenkins' report (AICPA, 1994) identifies investor's need for financial information, which was defined as including financial statements and related disclosures.

A recurrent feature of these publications is the emphasis placed on the *location* of the information. Accounting information is generally equated with information located in financial reports (ICAEW, 2009) with financial reports including annual reports, interim and preliminary reports (FRC, 2012). In general *financial*

statements seem to be a widely used referent for the definition of accounting information. This is perhaps a reflection of the fact that financial statements are the "end results of an accounting recordkeeping process that records the economic activities of a company" (CFA, 2008; p.12). Nevertheless, as emphasised in Beattie *et al*. (2007), it is important to clearly distinguish between types of information and the source document from which it is derived or the channel of communication of such information.

For the purpose of this research, accounting information is not limited to financial statements information only. Similar to Barker and Imam (2008), a more comprehensive definition of accounting information is adopted and comprises financial statement information items, their sub-components and information items that incorporates financial statement data in the form of ratios, valuation models or forecasts. This is illustrated in Figure 3.

Figure 3 : Definition of accounting information



Information generally contained in financial statements is considered accounting information. Conceptually, the information contained in financial statements is

defined by accounting conceptual frameworks. For instance, the IASB conceptual framework (2010) states that information contained in the financial reports are information about an entity's economic resources, claims against the reporting entity and the effects of transactions and other events that change those resources. In other words, it is information about a firm's financial position and performance, which are measured and summarised by specific elements in the financial statements. IASB (2010, par. 4.2) identify 'assets, liabilities and equities' as financial statement elements that portray an entity's financial position, while 'income and expenses' delineate the financial performance.

Further, reference to financial statement information need not be limited to the line items presented in the financial statements. The disaggregated components of asset, liability, equity, income or expenses are also considered as accounting information independent of its presence as a line item in the financial statements. For example, Selling General and Administrative costs are generally an item found in the income statements. However, these costs are an aggregation of different direct and indirect costs such as marketing expenses. Mention of these sub-components of financial statement line items are also coded as accounting information. This allows a comprehensive definition of 'accounting information' which is not limited to a source document (such as the financial statements) or the method of presentation (summary or detailed) of information therein.

Finally, the uses of accounting information in the form of accounting ratios, valuation models and forecasts are considered accounting information. Such uses could be deterministic or not. An example of a deterministic use of accounting information is given by accounting valuation models and accounting ratios. In these cases, accounting information serves as a direct input into well-defined models. For instance the Dividend Discount Model for firm valuation, which equates firm value to the discounted dividends as follows: $P = \frac{Div}{r}$[27], requires information about the firms' dividends which is an accounting information. Similarly, the use of accounting ratios also involves relating different types of accounting information in deterministic models. These could include only accounting information such as the ROE ratio given as

[27] Where P= price, Div = Dividends and r = discount rate.

$\frac{Net\ profit\ available\ to\ \ shareholders}{Shareholder's\ capital}$ or a combination of accounting and non-accounting information such as the $Price/_{Earnings}$ ratio, which combines accounting information (earnings) with non-accounting information (price). Historical accounting information could also be used in forecasting future accounting information such as earnings. However, unlike valuation models and ratios, the use of historical accounting information for forecasts is not deterministic (Abhayawansa *et al.* 2015). These are also included in the definition of accounting information, independent of the explicit or implicit method of forecasting.

5.4.2.2 Categories of accounting information

Having presented a definition of accounting information in the previous section, the categories of accounting information to be used in the proposed research is discussed in this section. First, the approaches to obtaining categories are discussed, second, the approach used in this research presented and justified and third, the categorisation scheme derived is presented and discussed.

Boyatzis (1998) identified 3 ways to develop thematic code/categories for the research, each with its benefits and limitations:

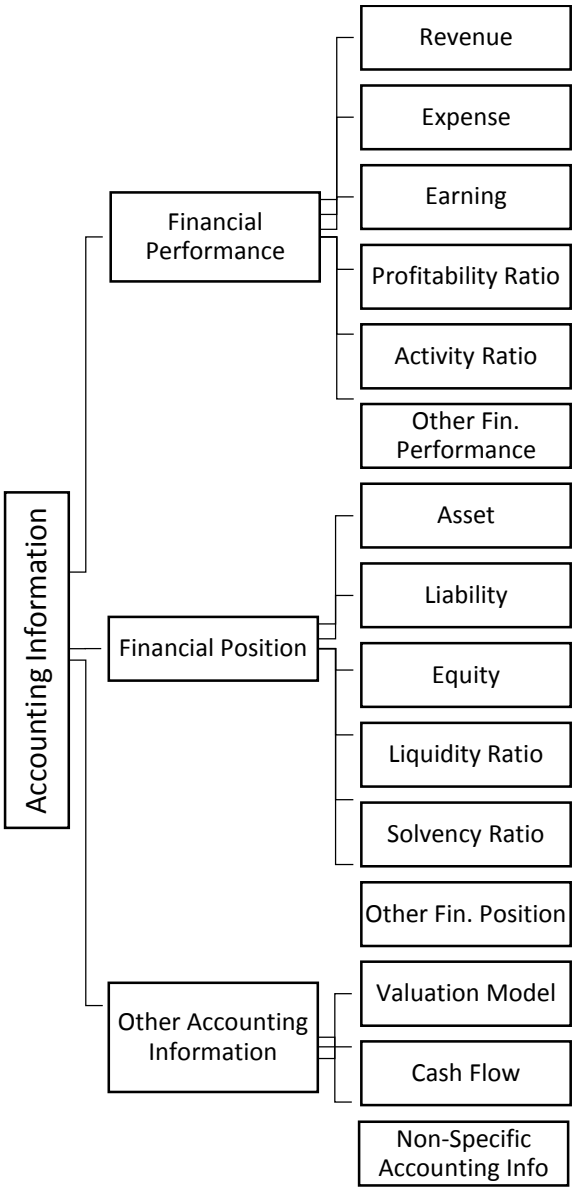
Prior Research Driven: This involves relying on categorisation schemes used in previous studies. The main benefit is the possibility of comparing results across studies, with findings being cumulative (Holsti, 1969). A major challenge of this approach is the difficulty in finding a suitable scheme as these will generally be driven by different research objectives. With particular consideration of studies that examine the content of analysts' reports, no comprehensive scheme is currently in place. Bricker *et al.* (1995) and Barker and Imam (2008) investigated the concept of earnings' quality and derived a list of possible keywords. Nevertheless, the list is not exhaustive and was principally tailored to suit the earnings' quality theme being explored. Breton and Taffler (2001) categorises accounting information into financial position and financial performance, including a few keywords into each of these categories. This was also found to be non-exhaustive. In essence, categories developed from other research projects are either too specific to the research question being investigated or are not exhaustive which violates a key principle of category construction.

Data driven research (Inductive): Categories are constructed based on the information contained in the data.

Theory driven: This involves developing categories which are consistent with a particular theory of interest.

For the purpose of this research, categories of accounting information were obtained through a combination of the three approaches. The IASB conceptual framework formed the basis for initial categorisation. It specifies the basic elements of the financial statements which reflect information about an entity's financial position and performance. The basic elements are Revenue, Expenses, Earnings, Equity, Assets and Liabilities. However, several types of information included as accounting information in the definition adopted for this study are not included in these items such as accounting ratios and valuation models. Hence, further reference was made to financial statements analysis textbooks, CFA curriculum study books, previous research papers on the content of analysts' reports (e.g. Breton and Taffler, 2001) and a pilot sample of analysts' reports. The resulting categories of accounting information used for this study presented in Figure 4:

Figure 4: Categories of accounting information



The figure presents the main categories identified for this study and was constructed to ensure the categories were exhaustive, independent, and mutually exclusive[28] and a single classification scheme was used to ensure maximum precision during coding. Though the main topic of investigation is accounting information, sub-categories of accounting information were derived to distinguish

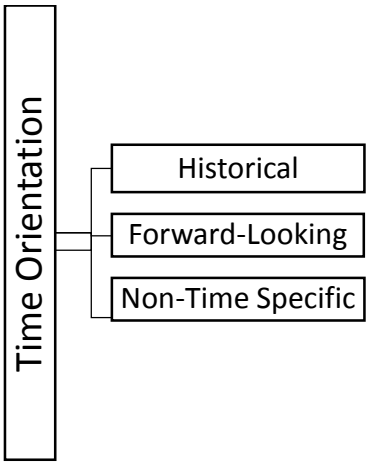
[28] The categories were derived principally from IASB conceptual framework on the basic elements of financial statements. Other alternative classification schemes were also considered. For instance, a classification scheme could be based on the reports that make up the financial statements (Balance sheet, Income statement, Cash flow statement etc). However, this approach doesn't ensure adherence to the principle of mutual exclusiveness as content of the income statements could appear in the balance sheet (e.g. retained earnings) or cash flow statement (e.g. operating expenses).

between financial position and financial performance related information, consistent with IASB framework discussed above and previous studies such as Breton and Taffler (2001). Financial position category includes information relating to assets, liabilities, equities, liquidity and solvency ratios and for completion, a category for other financial position related information not classifiable in the other categories. Similarly, the financial performance category includes information relating to revenues, expenses, earnings, profitability and activity ratios and any other financial performance related information not classifiable in any of the other categories. To ensure that the categories are exhaustive, consistent with the principles set forth in Holsti (1969), an additional category was derived to capture accounting information not classifiable in the other categories. This include information about cash flows, accounting valuation models and accounting information that is not specific enough to enable classification in any of the previous categories. The division into lower level categories enables richness of data and provide additional details to enhance understanding of types of accounting information used by analysts, which is an important extension to the current literature on analysts' reports.

Beyond the main topic of interest, there are several qualitative attributes of the information that could be further explored. Beattie *et al.* (2004) highlighted the importance of capturing multiple facet of disclosure and introduced a holistic disclosure profile which explored other dimension of corporate disclosure content including the time dimension, the financial/non-financial, quantitative/qualitative etc. Given the aim off the present study is to explore the variation in the extent of use of accounting information. The choice of additional qualitative category is influenced by this context and research objective. Of the several disclosure dimensions explored in extant literature (such as qualitative/quantitative, financial/nonfinancial, monetary/non-monetary, tone, location, etc.), the most relevant attribute which is likely to yield additional insight into the relevance of accounting information to analysts, is the time dimension. This allows the differential role of accounting information to be further explored. Both the stewardship and valuation role of accounting information has been acknowledged in accounting literature. The stewardship role is mostly concerned with analysis of (historical) managerial performance, while valuation generally focuses on assessment of future prospects. Consequently, a

distinction based on time dimension is useful in making inference regarding the rationale of use of accounting information by analysts. The two main time dimension categories are 'historical' and 'forward looking'. However, to ensure that the category is exhaustive, a third category which captures non-time specific disclosures is included as in Beattie *et al*. (2004). The resulting classification scheme for the time dimension is as follows:

Figure 5: Time-orientation category



5.4.2.3 Developing the coding instrument

Having identified the recording and measuring units and categories of accounting information, a detailed coding instrument was developed which provides guidelines on how to recognise information belonging to difffernt categories. To enhance reliability and validity of coding, Boyatzis (1998) suggested that the coding rules should include the following items for each category:

(i) A label
(ii) Definition of the category
(iii) Description of how to flag the category
(iv) Any qualifications or exclusions from the category
(v) Examples

In line with these suggestions and consistent with the coding instruments developed in previous research such as Beattie *et al*. (2004) and Abhayawansa (2010), a detailed coding instrument was developed which contained each of the following items for each category of accounting information and time dimension. The coding rules are contained in appendix B and provides detailed information on classification of text portions into the different categories. The coding instrument was developed over several months with references made to IASB conceptual framework, accounting texts, the CFA examination study materials and previous studies of the content of analysts' reports (e.g. Bricker *et al*. 1995; Previts *et al*. 1994; Rogers and Grant 1997; Abdolmohammadi *et al*. 2006 and Nielsen 2007) and through test coding of a sub-sample of reports. Analysis of a pilot sample is an essential step in the development and application of a coding scheme (Weber, 1990). For this purpose, a sample of 21 firms, stratified by industry (at the 2 digit SIC code level) were selected. This enabled the differences in terminologies used by analysts across industries to be captured and included in the coding rules. Text was coded and presented for discussion with supervisors. Following discussion, the coding instrument was revised and extended. Examples of issued discussed during the test coding stage and coding rules agreed on, which were incorporated into the final coding instrument are presented in Table 6

Table 6: Examples of coding decisions

Cases	Illustrative examples	Coding principles
Use of words such as 'guidance' or 'estimates'	TJX provided detailed 1Q and 2011 guidance.	Code only when accompanied by an underlying accounting keyword, e.g. references to earnings guidance
Use of accounting keywords to qualify strategies, programs or activities	Under the current $8 billion share buyback program	Code only if text portion is contain the value of the program
References to 'acquisition'	The acquisition closed in August 2010	Code only if accompanied by related accounting information and their value, e.g. goodwill

5.4.2.4 Identifying text units

As previously mentioned in section 5.4.1.5, text units are used as recording units for this part of the thesis. Unlike other units discussed in that section, text units are not grammatical units with clearly distinguished borders. Hence, during the test coding stages, decisions were made on how to split text portions into units for coding. Choices were made in light of the definition of text units adopted for this research, which is a text portion containing "a single piece of information that is meaningful in its own right" (Beattie *et al*. 2004). In several cases, sentences contain only a single piece of accounting information, codable to a single accounting and time category. For example:

> We are raising our 2011 ZMH adjusted EPS estimate $0.03 to $4.73.
>
> *(Source: Northcoast Research on Zimmer Holdings, 28.01.11)*

In such a case, the sentence constitutes the text unit and is coded to the appropriate accounting and time category. There were also cases of sentences with two or more types of accounting information. Hence sentences were split into different portions for coding. During the test coding stages, it emerged that there were two main types of such cases. The first is a situation in which sentences contain multiple accounting information or non-accounting information but are discussed in separate and clearly identifiable clauses, often separated by punctuation, conjunctions or causal words. For example the extract below:

> International operating margins decreased 80 basis points to 13.0%, due to higher coffee prices.
>
> *(Source: Williams Capital Group on Starbucks corporations, 8.11.11)*

In the above sentence, the analysts refer to operating margins - an accounting information and coffee prices – a non-accounting information. However, both pieces of information are discussed in clauses separated by a comma and the causal words "due to". In such cases, each clause relating to accounting information constitutes a text unit and is coded into the related accounting and time category.

The second case is a situation in which sentences involves a conflation of accounting information not clearly distinguished by separate clauses. These are cases in which two or more accounting information is intertwined in a sentence. For example:

> We are increasing our FY2012 revenue and EPS estimates to $2.271 billion and $3.28 from $2.204 billion and $3.22, respectively.
>
> *(Source: Credit Suisse report on BMC Software Inc., 5.5.11)*

In the above statement, analysts refer to both revenue and earnings. Although both items relate to financial performance and are components of the income statement, the sub-level categorisation adopted in this study, necessitates coding into different categories. This raises the question of the most reliable and valid procedure for splitting such text portions into appropriate categories. First, reference was made to existing content analytic studies. Abhayawansa (2011) suggest double counting of the portions of text common to both items. In other words, the sentence would be classified twice; first to a revenue category and second to an earnings category. However, this results in text portions classified in more than one category and word counts greater than those used by analysts, which gives a false impression of relevance (as measured based on number of words). However, the decision was made earlier on in the research to uphold mutual exclusive units such that percentages would sum to 100%. Hence, this approach was not adopted.

A second approach is that used by Smith and Taffler (2000) which consists in splitting sentences equally into different categories. For instance, they suggested that if four themes were discussed in a sentence, then each category should be assigned a score of 0.25 of the sentence. Hence, only a proportion of the sentence is assigned to the category. For instance, in the above extract, there are 20 words in total. Under this approach, this will require an equal split of the sentence such that the revenue category is assigned 10 words and the earnings category 10 words. This option results in a reliable split of the statement as it is based on word count. Hence, it was adopted in this study. This approach also has the added advantage of overcoming the limitation of double counting. However, it is more laborious as the common words have to be counted and then split between information items before coding into appropriate categories.

Table 7: summarises the cases discussed and the decisions made about coding

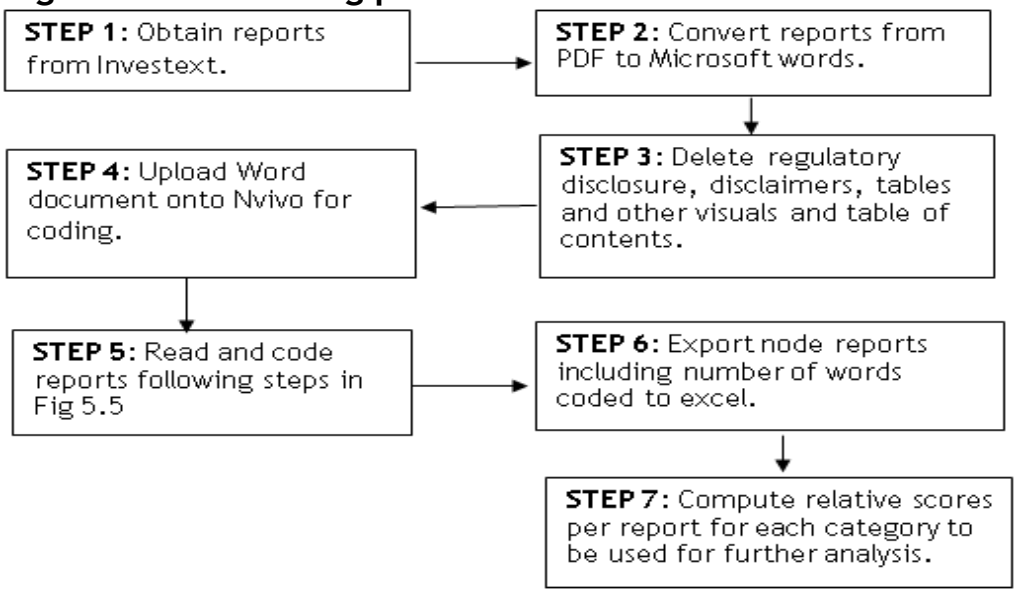
Cases	Coding principles
Sentence contains single piece of information	Code entire sentence to the accounting and time category
Sentence contains more than a piece of information but these are contained in separate clauses	For each clause containing accounting information, code all words in the clause to the related accounting and time category
Sentence contains more than a piece of information but these are intertwined within the sentence or clauses	Split[29] the sentences equally between the information types and code portions with accounting information to appropriate accounting and time categories.

[29] In a situation in which common words can't be equally split between text portions (e.g. odd numbers to be split among even text portions), the assigning of words to text portion is to begin with the first text portion until extra words have been fully assigned. The alternative is to split the words proportionally. However, for practical reasons, this could not be achieved. It is not envisaged that an extra word missing from a text portion would affect results significantly. Moreover, the assignation is random as this only occurs in such cases of conflation in which odd number of words are to be split among even text portions.

5.4.2.5 The coding process

The data collection process involved several steps which is illustrated in Figure 6 and discussed below:

Figure 6: The coding process

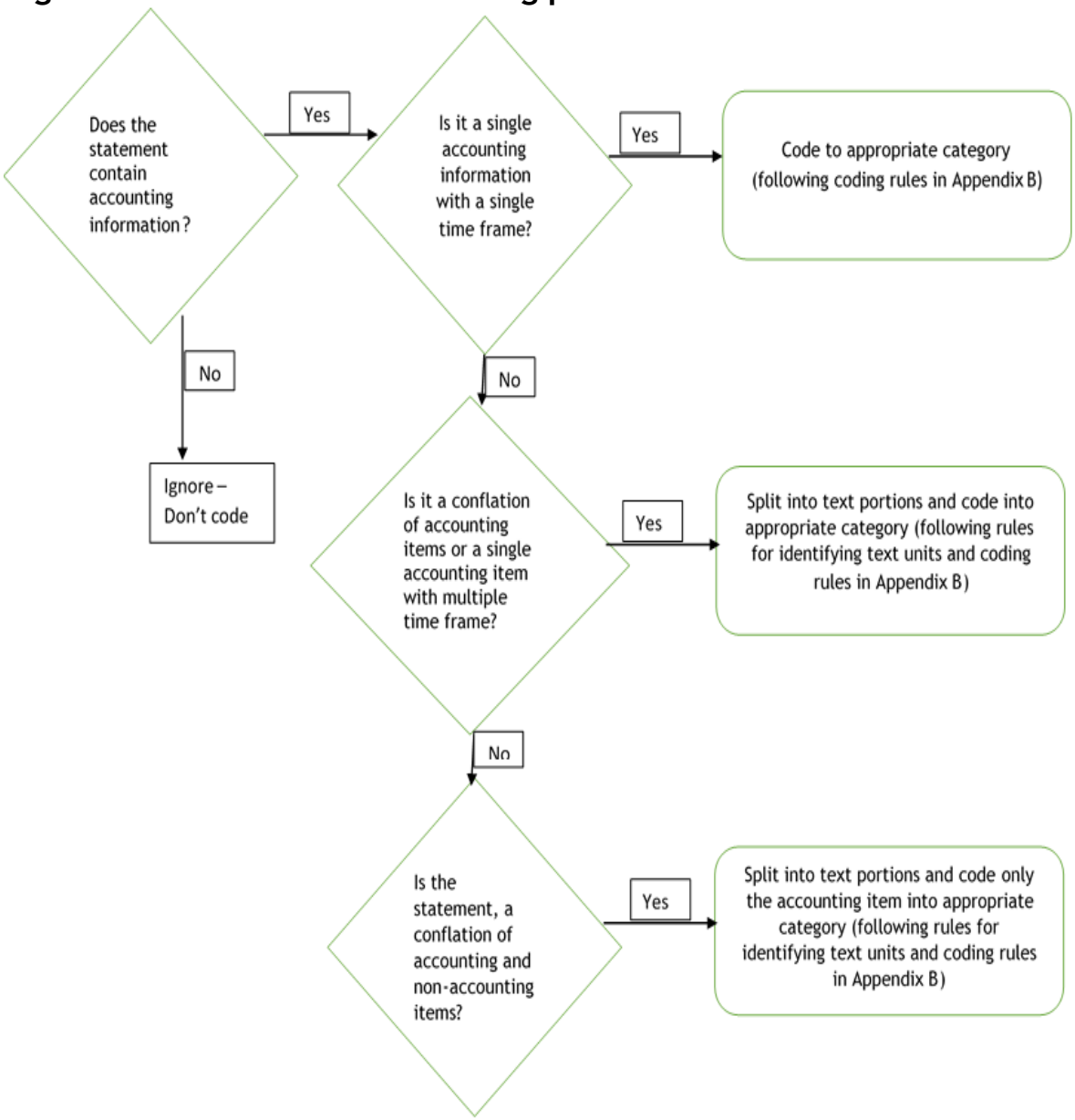


Analysts' reports downloaded from *Investext* were in PDF format. Thus, the first step was to convert these into word documents for upload on to Nvivo. This process mostly consisted of copying the text from the PDF documents into a word document. Only the main body of narratives in the reports were coded. Hence, information in the appendix such as regulatory disclosures, disclaimers and analysts' certification, distribution of recommendation, tables and figures were not copied. The converted reports were then uploaded onto Nvivo 9, the computer program used to code text portions into various categories.

Nvivo is a computer program used for analysis of non-numerical and unstructured data. It is generally used for qualitative research and possess features useful for coding and searching patterns of coding. In Nvivo workspace, users are able to create categories for coding called nodes which could be unstructured or hierarchical. Coding is done by linking the portions of text to the nodes which represent them. Following coding, Nvivo possess several features which enable users to analyse, organise and extract relevant information from the documents. It has been used in previous studies of the content of analysts' reports (Abhayawansa 2010) and corporate reports (Beattie *et al*. 2004).

For coding purposes, the hierarchical structure was used to create nodes in Nvivo. Two main categories were created to reflect the two-dimensional structure of the coding, i.e. the accounting topic and the time orientation of the statements containing accounting information. Lower level categories were also created to reflect the classifications presented in Figures 4 and 5. Following upload of analysts' reports onto Nvivo, each report was read and references to accounting information were coded into relevant accounting and time categories. Figure 7 present a flow chart of the coding process in Nvivo.

Figure 7: Flowchart of the coding process



As illustrated, the process begins with reading a sentence. If the sentence contains information about a single accounting information, this was coded to the relevant category. If the sentence contains more than one accounting information, it was split into portions that were coded into the appropriate categories. Splitting of sentences were made in light of the discussion presented in section 5.4.2.5 on identifying text units.

Following coding, the query function in Nvivo was used to extract the results of the coding. For each report, the total number of words coded to each of the accounting and time categories were extracted. Checks were made to correct errors during coding. For instance, the total word count must be equal for the accounting and time categories and where there were differences, these were checked and corrected. Subsequently, the coding results were exported to Microsoft Excel and the content scores calculated for each category.

5.4.3 Reliability and validity

One of the commonly cited limitations of content analysis is the subjectivity in the interpretations of coding rules which poses a threat to the reliability and validity of the study (Clatworthy and Jones 2003, Linsley and Shrives 2006). Reliability and validity are interrelated concepts as Morris (1994, p.905) suggests that "if valid inferences about the symbolic content of the message are to be drawn, the content analysis classification scheme must be reliable in terms of consistency and reproducibility". Thus reliability is a necessary but not sufficient condition for validity, as "a measure cannot be valid if it is not reliable" (Neuendorf, 2002, p.113). This section explains types of reliability and validity and the steps taken as part of this research to achieve these.

5.4.3.1 Reliability

For content analysis, Neuendorf (2002, p.112) defines reliability as the "extent to which a measuring procedure yields the same results on repeated trials." The notion is similar to the concept of replicability in Seale (1999). Krippendorff (2004) identifies three types of reliability in content analysis:

(i) Stability measures the degree to which a coding process produces the same results on repeated trials over time. This is tested through the re-coding of the same text by the same coder after some time has elapsed. The aim is to test the consistency of the coder over time, which may be threatened by distractions or carelessness, unclear coding rules or tiredness. Moreover, Krippendorff (2004) argues that "human characteristics of learning through practice" may improve performance over time such that there could be disagreements in coding over time. Stability test is a first step but is not sufficient in measuring the reliability of the coding process.

(ii) Reproducibility measures the extent to which different coders, working independently, consistently apply same coding instructions in classifying the same units. The aim is to test inter-coder differences and this is perceived as a stronger measure of reliability than stability (Krippendorff 2004).

(iii) Accuracy measures the extent to which a coding process conforms to a standard (Krippendorff 2004). This implies comparing the results of the coding process to a process that is taken to be correct in principle.

In general, reliability in content analysis is influenced by various factors such as clarity of the categories and the coding rules, degree of ambiguity in the data and coders' skills (Holsti, 1969; Morris 1994). Thus, several steps were taken to ensure the reliability of the coding process in this study. First, the coding instrument was well-specified as categories were developed with reference to relevant texts such as the IASB framework. This ensured that widely shared meaning of what constitutes accounting information and lower level categorisation of accounting information were reflected in the coding scheme. Second, categories of accounting information were developed based on the principles advocated in Holsti (1969) and previously discussed in section 5.4.2.1. Principles of mutual exclusive and exhaustive categories are particularly useful in reducing ambiguity during the coding process. Third, coding instructions were detailed and adhered to principles advocated in Boyatzis (1998) which suggest the inclusion of clear flags, qualification and relevant examples. Unlike other types of information such as risk disclosures or intellectual capital information that could be expressed using a variety of terminologies, which are open to interpretation (Linsley and Shrives, 2006), accounting has its own language which is easily identifiably by certain key words such as expense, revenue, depreciation etc., which further enhanced the reliability of the coding process. However, analysts were found to use several abbreviation and acronyms which may threaten the reliability. Hence, to further eliminate ambiguities, such abbreviations and terminologies were clearly defined and further incorporated into the coding instrument. The list of such terms is contained in Appendix B.

To formally assess the reliability of the coding process, test of stability and reproducibility were conducted. The first case involved testing the consistency of the main coder over a space of two months while the second test involved the use of a second coder to check the level of agreement with the main coder. Following Beattie *et al*. (2004), blocks of texts (5-10 sentences each) were extracted from a randomly selected sample of 29 reports. These were

extracted from different sections of the reports and were coded on two occasions by the main researcher and a sub-set was coded independently by the main researcher and a second coder with a non-business background[30]. Following coding, the levels of agreements were calculated.

There are several measures of inter-coder and intra-coder agreements used in content-analytic studies. The simplest and most intuitive is the percentage of agreement, which is the ratio of coding agreements over the total number of coding decisions made. However, this has been largely criticised in existing content analysis literature due to its failure to account for agreement by chance among coders (Milne and Adler, 1999; Krippendorff, 2004). As agreement by chance increases as the number of categories decreases (Holsti, 1969), the measure is biased in favour of coding dimensions with fewer categories such as the time-dimension in this research. In response to these criticisms, several alternative measures have been developed to measure the extent of agreement between two coders, which adjusts for the extent of agreement by chance, such as Krippendorff's alpha(α), Scott's pi (π) and Cohen's kappa (κ) (Neuendorf, 2002). Following prior content analytic studies of analysts' reports (e.g. Fogarty and Rogers 2005; Barker and Imam 2008 and Abhayawansa 2010), Cohen's kappa (κ) is used to measure reliability in this research. This was computed using SPSS and is given by:

$$\kappa^{31} = \frac{(\%\ observed\ agreement - \%\ expected\ agreement)}{(1 - \%\ expected\ agreement)}$$

[30] To enhance reliability, coding was preceded by a period of pre-testing and coder training. Holsti (1969) advocates this as a means of ensuring that coders rely on the same aspects of their experience in their coding decisions and it enables coders to familiarise with the texts, categories and related coding instruction (Abhayawansa 2011). The training consisted of discussing the research objective, analysts' reports, and coding instruction with the second coder. This was followed by an illustrative example by the main coder and subsequently a check of the second coder's understanding of the requirements of the task through test coding of a block of text. There were cases in which the coding instructions were misinterpreted or not understood. Hence, the coding instrument was further refined at this stage.

[31] This is similar to Scott's pi (which is the other widely used measure in accounting research), with the only difference being the measurement of % expected agreement.

Table 8 and 9 show the results of the stability and reproducibility tests respectively.

Table 8: Reliability measures for test of stability

	Percent Agreement	Cohen's Kappa	Agreements	Disagreements	Total Units
Sentences	96%	92%	279	11	290
Accounting	97%	94%	174	6	180
Time	92%	86%	138	12	150

Table 9: Reliability measures for test of reproducibility

	Percent Agreement	Cohen's Kappa	Agreements	Disagreements	Total Units
Sentences	98%	96%	192	4	196
Accounting	87%	83%	123	27	150
Time-first	82%	64%	96	21	117
Time	96%	92%	107	4	111

The tables show the percent agreement, Cohen's kappa, number of agreements, number of disagreements and total decision units. In the first instance, for each sentence which form part of the block of text extracted from analysts' reports, decision was made whether it contains accounting information or not. Second, for the units containing accounting information, decision was made as to the category of accounting information. This level of coding involved most categories and expected agreement by chance decreases as the number of categories increases, thus the adjustment for agreement by chance is lower for this level of coding (resulting in smaller differences between percentage agreement and kappa). Third level of coding is decision with regards to the time orientation of the accounting statements[32]. The percentage agreement in the first round of coding

[32] The same blocks of text were coded for accounting and time orientation. However, the difference between total units for the topic (accounting) dimension and the time dimension (150 cf 117 in Table 9) stemmed from the fact that a sentence could contain two or more categories of accounting information but are all historical references. This results in multiple decisions for the accounting dimension but only one decision for the time dimension. This was the general pattern in analysts' reports, hence there were more decisions for the topic category than the time category.

was 82%, which is similar to the topic dimension. However, as the time dimension contained lesser categories, Cohen's kappa was 64%. Following discussion and further clarification, a re-coding was carried out with more acceptable level of reliability[33].

While there is no theoretically agreed standard when measuring reliability, generally accepted lower limits range between 70% (Boyatzis 1998) and 80% (Krippendorff 2004). The percentages reported in Tables 5.7 and 5.8 compare well with these levels as well as those reported in other studies. For instance, Barker and Imam (2008) reported a measure of Cohen's kappa of 70%, while Abhayawansa (2010) reported values of Cohen's kappa between 79% and 97%. These are also similar with Scott's pi reported in other studies of accounting narratives such as 70%-84% in Beattie *et al.* (2004) and 83%-87% in Abraham and Cox (2007).

5.4.3.2 Validity

In addition to reliability, coding schemes must also be valid (Morris, 1994). Validity concerns the extent to which the coding procedure measures the intended concept (Holsti 1969, Neuendorf 2002, Krippendorff 2004, and Abhayawansa 2010). In general, different types of validity are discussed in existing literature. For instance, Brennan *et al.* (2009) discusses types of validity listed in Morris (1994), which are:

(i) Construct validity: This is the extent to which the variables are correlated with measure of the same construct

(ii) Hypothesis validity: This is the extent to which the variables behave as they are supposed to with regards to other variables

(iii) Face or content validity: This is the extent to which the classification scheme appears to measure the intended construct.

[33] Much of the differences in the first round were mostly cases of oversight by the second coder and coding instructions were not amended at this stage.

(iv) Semantic validity: This is the extent to which persons familiar with the language and texts agree that the list of words placed in the category have similar meanings.

In addition to these, Holsti (1969) also mention concurrent and predictive validity while Krippendorff (2004) refer to sampling validity, structural validity, functional validity, correlative validity and criterion validity. In general, establishing validity in content analysis concerns choices made with regards to the overall research design, choice of categories, choice of recording units, coding instructions, sampling methods and reliability (Holsti, 1969; Abhayawansa 2010). The choices made have been previously discussed and are restated briefly with the scope of highlighting how they helped establish the validity of this study.

First, to establish validity, the sampling process must ensure the sample is representative of the population from which it is obtained. This is referred to as sampling validity in Krippendorff (2004) and external validity in Neuendorf (2002). Generally, this is achieved through random selection of the samples used in the research. To address the research question in this study, a matched sample of reports was necessary. However, bias introduced through matching were controlled for by the use of conditional logistic regression as advocated by Cram *et al.* (2009) as a means of obtaining more valid analysis for matched samples.

Second, in developing the categorisation scheme, accounting information was conceptualised with reference to the IASB conceptual framework, which is an authoritative text for accounting practice. Moreover, categories of accounting information were also constructed with reference to existing studies (such as Breton and Taffler, 2001) and accounting texts, including the CFA exam study material. Hence, the classification scheme used in this study is based on the perception of accounting information as generally defined in practice and academic community, thereby enhancing content/face and semantic validity. Morris (1994) states that the level of agreement between coders can be a surrogate of semantic validity. As discussed in the previous section, tests of reliability generally produced acceptable levels of reliability and support the semantic validity of the coding process.

Third, the choice of recording and measuring units as discussed in section 5.4.1.5 ensures that the content score which measures the extent of use of accounting information in this study captures the words used to discuss accounting information in analysts' reports and mutual exclusiveness of the units ensure that the measure reflects only accounting information in the related category.

5.5 Chapter summary

This chapter was aimed at describing the content analysis method used to generate the dependent variable for the first part of this thesis.

First, the approaches to content analysis were reviewed and their merits for the present research considered. It was stated, that to address the first research objective, a thematic, volumetric, manifest and manual content analysis approach is used.

Second, research design choices pertaining to the data used in the research were discussed including data sources, types of analysts' reports, companies and analysts-firms. The final sample consist of 288 results reports issued on 144 S&P 500 companies and issued by analysts employed by investment banking and independent research firms.

Third, the content analysis process was discussed in detail. Following a review of the merits, of different recording and measuring units used in existing content analytic studies, text units and word counts were selected for this research. Also, a two-dimension categorisation scheme was developed to conduct the content analysis. This included a main topic category for capturing accounting information and a time category for capturing the time orientation of the accounting statements. The categorisation scheme was developed with reference to existing literature and included a definition of accounting information to include information generally presented in the financial statements, its sub-components and uses of such information in the form of forecasts, ratios and accounting valuation models.

Finally, the chapter discussed issues relating to reliability and validity of the coding process. Steps taken to achieve these and results of inter-coder and intra-

coder reliability were reported based on the percentage agreement and Cohen's kappa. The agreement level achieved was comparable to other studies and in line with generally accepted levels.

CHAPTER 6: Use of accounting information in analysts' reports: empirical results and discussion

6.1 Introduction

This chapter reports and discusses the results of the empirical analysis relating to the first objective of the thesis which is to examine the factors that explain variation in use of accounting information by analysts through content analysis of their reports. The measures from the content analysis process are the dependent variables for this study and tests of association between these and a range of company, analysts and report specific attributes are conducted. These are reported and discussed in this chapter.

The chapter is structured as follows: section 6.2 provides descriptive statistics of the independent and dependent variables used in the study. Section 6.3 presents and discusses the results of univariate analysis. Section 6.4 presents and discusses the results of multivariate analysis testing the association between the extent of use of accounting information and the independent variables. This is followed by discussion of robustness tests in section 6.5 and finally 6.6 concludes the chapter.

6.2 Descriptive statistics

This section presents and discusses descriptive statistics for the variables used in this study. Results are presented in two separate sections, section 6.2.1 presents the results for the independent variables and section 6.2.2 presents the results for the dependent variables.

6.2.1 Independent variables

The descriptive statistics for the company, analysts and report features which constitute the independent variables for this study are presented in Table 10, 11, 12 and 13.

Table 10: Descriptive statistics – company-specific variables

Panel A: Continuous variables

	N	Mean	Median	Std. Deviation	Min	Max
ΔP_EPS	280	0.69	0.23	1.88	-2.69	14.91
Risk	280	0.04	0.03	0.03	0	0.17
Growth	274	0.04	0.04	0.10	-0.31	0.54
Intangibles	280	0.23	0.19	0.19	0	0.69
MTB	268	3.49	2.78	2.36	0.80	14.59
Size	280	9.68	9.52	1.12	7.38	12.47
Leverage	278	0.19	0.19	0.14	0	0.61

Panel B: Categorical variables

Industry	N	Percent
HTC	114	40%
LTC	174	60%
Total	**288**	**100%**

This table shows the summary statistics for company specific characteristics. ΔP_EPS is the percentage change in EPS from prior year, Risk is the standard deviation of EBIT over five years (2006-2010) scaled by average assets over the time period. Growth is the cumulative annual growth rate in sales per share over the previous five years from 2006-2010. Intangibles is the proportion of intangible assets over total assets. MTB is the ratio of market value of equity to book value of equity for the fiscal year end 2010. Size is the natural logarithm of market value for the fiscal year end 2010. Leverage is the long term debt to total asset ratio. All variables were obtained from Computstat. Companies are classified into two groups, i.e. hi-tech (HTC) and low-tech (LTC). Classification of companies into HTC and LTC is based on the classification scheme used in Francis and Schipper (1999).

Table 11: Descriptive statistics: analyst-specific and report-specific variables

Analysts' Characteristics

CFA	N	Percent
Non-CFA	154	55%
CFA	125	45%
Total	**279**	**100%**

The table shows the frequency for the analysts and report attributes. CFA is a dummy variable which takes the value of 1 for reports written by at least a CFA qualified analysts and 0 if none of the authors possess the CFA qualification.

Table 12: Descriptive statistics by industry group

	HTC	LTC	Diff
ΔP_EPS	0.56	0.78	-0.22
Risk	0.04	0.04	0
Growth	0.07	0.02	0.05***
Intangibles	0.28	0.19	0.09***
MTB	6.13	4.87	1.26
Size	9.9	9.54	0.36*
Leverage	0.2	0.2	-0.02

This table shows the mean for firm specific characteristics across the two industry groups. ΔP_EPS is the percentage change in EPS from prior year and one year ahead, Risk is the standard deviation of EBIT over five years (2006-2010) scaled by average assets over the time period. Growth is the cumulative annual growth rate in sales per share over the previous five years from 2006-2010. Intangibles is the proportion of intangible assets over total assets. Market-book is the ratio of market value of equity to book value of equity for the fiscal year end 2010. Size is the natural logarithm of market value for the fiscal year end 2010. Leverage is the long term debt to total asset ratio. All variables were obtained from Computstat.

Table 10 shows that the sampled companies have improving profits as the mean (median) percentage change in EPS from prior year is 69% (23%).

The sampled companies appear to have a low risk with a mean (median) value of 0.04 (0.03). The mean (median) 5 year sales growth rate is 4% (3%). Additionally, intangible assets as a proportion of total assets have a mean (median) value of 23% (19%). Panel B also shows that 40% of the companies in the sample are high-tech companies, while 60% are low-tech. Table 12 shows the mean values for the independent variables across these two industry groups. As would be expected, the mean values for sales growth and intangibles are significantly higher for high-tech companies compared to low-tech companies. Thus, part of the variation in these variables can be explained by industry affiliations.The mean (median) market to book ratio is 3.49 (2.78) while the mean size, as proxied by the natural logarithm of the companies' market value, is 9.68 and ranges from a minimum of 7.38 to 12.47. Finally, the descriptive statistics reveal that the companies are moderately levered. The mean (median) leverage value, measured by the long-term debt to equity ratio, is 19% (19%) and ranges from 0 to 61%. This is consistent with previous studies. For instance, for a large sample of US firms, Kothari *et al.*

(2009) reported a mean size of 6.1 and a mean leverage value of 16.9%. The results in table 11 show that consistent with prior studies (e.g. Asquith *et al*. 2005, Abhayawansa and Guthrie, 2012, and Twedt and Rees, 2012), analysts are reluctant to issue sell recommendations with only 7% of the sample in this category. Overall, positive recommendations are proportionately higher across the sampled companies with a 52%, closely followed by hold recommendations with a 41%, indicating a tendency towards optimism by analysts. Table 13 presents the mean values of the independent variables across recommendation types.

Table 13: Descriptive statistics by types of recommendations

Panel A: Descriptive Statistics by recommendation

Recommendation	Positive	Neutral/Negative	Difference
ΔP_EPS	0.8	0.58	0.22
Risk	0.04	0.03	0.01
Growth	0.04	0.03	0
Intangibles	0.22	0.24	-0.03
MTB	4.69	6.15	-1.46
Size	9.75	9.55	0.21
Leverage	0.19	0.21	-0.02

Panel B: Recommendation types

Recommendation	N	Percent	Recommendation
Negative	21	7%	Negative
Neutral	116	41%	Neutral
Positive	147	52%	Positive
Total	**284**	**100%**	**Total**

The table presents the mean for the company attributes across recommendation types. ΔP_EPS is the percentage change in EPS from prior year and one year ahead, Risk is the standard deviation of EBIT over five years (2006-2010) scaled by average assets over the time period. Growth is the cumulative annual growth rate in sales per share over the previous five years from 2006-2010. Intangibles is the proportion of intangible assets over total assets. Market-book is the ratio of market value of equity to book value of equity for the fiscal year end 2010. Size is the natural logarithm of market value for the fiscal year end 2010. Leverage is the long term debt to total asset ratio. All variables are obtained from Computstat. Positive is for positive recommendations such as buy/strong buy, outperform etc. Neutral is for neutral recommendations such as hold, equal weight, and market perform while negative is for negative recommendations such as sell, underperform etc.

The results show that, compared to negative and neutral recommendations, firms with positive recommendations have larger changes in EPS from the prior year with a mean value of 80% compared with the mean of 58% for the negative and neutral sub-sample. Firms with positive recommendations are also larger, have lower leverage levels, higher volatility of earnings, higher sales growth and lower proportion of intangible assets and market-to-book ratio. Nevertheless, the differences between the sub-samples are not statistically significant.

6.2.2 Dependent variables

The dependent variables measure the extent of use of accounting information in analysts' reports. As previously discussed in Chapter 5, there are two main categories: the accounting information category and the time-orientation category. The descriptive statistics for the accounting category is presented in Table 14, while the descriptive statistics for the joint coding of accounting information and the time orientation are presented in Table 15. Panel A of both tables present the results for the top-level categories and panel B presents the results for lower level categories of accounting information.

Table 14: Descriptive statistics for the dependent variables

	N	Mean	Std. Deviation	Min	Max
Panel A: Level 1 categories					
FPER	288	29.77	14.47	2.45	72.94
FPOS	288	4.14	3.76	-	17.96
OACC	288	5.04	4.72	-	25.63
ACC	288	38.94	16.06	3.72	87.54
Panel B: Level 2 Categories					
Financial Performance					
ACT	288	0.07	0.35	-	3.05
EARN	288	8.67	6.46	-	29.80
EXP	288	4.03	3.95	-	23.83
OFPER	288	0.45	0.80	-	6.10
PRATIO	288	5.49	4.65	-	27.71
REV	288	11.06	8.38	-	46.57
Financial Position					
ASS	288	1.32	1.94	-	13.02
EQ	288	1.59	2.27	-	11.84
LIAB	288	0.54	0.88	-	4.88
LIQ	288	0.00	0.05	-	0.62
OFPOS	288	0.28	0.59	-	4.11
SOL	288	0.42	0.93	-	6.44
Other Accounting Information					
CF	288	1.42	2.06	-	14.02
NSA	288	0.41	0.91	-	8.41
VM	288	3.17	4.22	-	25.63

This table shows the summary statistics of the extent of use of accounting information. FPER – financial performance, FPOS – financial position, OACC – other accounting information, ACC – accounting information. ACT – activity ratios, EARN – earnings, EXP – expenses, OFPER – other financial performance, PRATIO – profitability ratios, REV – revenue, ASS – asset, EQ – equity, LIAB – liability, LIQ – liquidity ratios, OFPOS – other financial position, SOL – solvency ratios, CF – cash flow, NSA – non-specific accounting information, VM – Valuation models.

6.2.2.1 Accounting information

Panel A of Table 14 shows that accounting information is a relevant part of analysts' reports occupying on average 38.9% of the content of the reports. The extent of use of accounting information is also quite variable, ranging from a minimum of 3.7% to a maximum of 87.5%. The non-zero minimum value also reveals that accounting information is present in all of the sampled reports.

Within accounting information, financial performance information is mostly discussed, with a mean of 29.8% and ranges from 2.45% to 72.9%. This also reveals considerable variability in the use of financial performance information across all of the sampled reports. Financial position and other accounting information are less discussed in analysts' reports. The former has a mean of 4.14%, ranging from a minimum value of 0% to a maximum of 18%, while the latter has a mean of 5.04%, a minimum of zero and a maximum of 25.63%. These figures are relatively lower than those reported for the financial performance category and reveal that financial performance information is much more important to analysts than other types of accounting information and not all analysts' reports contain financial position and other accounting information.

Panel B presents the descriptive statistics for lower-level categories of accounting information categorised into financial performance, financial position and other accounting information.

Financial Performance

Panel B reports that revenue is the most frequently mentioned financial performance information item, on average accounting for 11.06% of the words in analysts' reports with a maximum of 46.57%. Following closely is information about earnings, with a mean reference of 8.67% words in analysts' reports and a maximum of 29.8%. Other sub-categories of financial performance are less referred to. For instance, profitability ratios on average occupy 5.49% of the content of analysts' reports, expenses occupy 4.03%, and activity ratios have a mean word count of 0.07% and other financial performance 0.45%.

Financial Position

Within the financial position category, the two most discussed types of accounting information are equity and assets, with a mean of 1.59% and 1.32%, respectively. This is closely followed by references to liabilities with a mean reference of 0.54%. Financial position ratios such as liquidity and solvency ratios are rarely discussed in analysts' reports. Overall, financial position-related information occupy less of analysts' reports.

Other Accounting Information

The category for other accounting information contains references to non-specific accounting information, cash flows and accounting-based valuation models. The latter two sub-categories occupy more of the content of analysts' reports with a mean of 1.42% and 3.17%, respectively.

6.2.2.2 Time orientation and accounting information

Table 15 presents the results of the joint coding of accounting topics and their time orientation. It reveals the proportions of each accounting information category that are forward-looking, historical or non-time specific.

Table 15: Descriptive statistics for the joint coding of accounting information and time

	Forward-Looking	Historical	Non-Time Specific	Total
Panel A: Level 1 categories				
FPER	11.16	16.37	2.25	29.77
FPOS	1.16	2.36	0.63	4.14
OACC	2.71	1.54	0.78	5.03
ACC	15.03	20.28	3.65	38.94
Panel B: Level 2 Categories				
Financial Performance				
ACT	0.01	0.06	0.00	0.07
EARN	4.33	4.04	0.31	8.67
EXP	1.44	2.11	0.48	4.03
OFPER	0.16	0.22	0.08	0.45
PRATIO	2.09	2.92	0.49	5.49
REV	3.14	7.04	0.89	11.06
Financial Position				
ASS	0.31	0.80	0.20	1.32
EQ	0.55	0.82	0.21	1.59
LIAB	0.13	0.33	0.08	0.54
LIQ	0.00	0.00	0.00	0.00
OFPOS	0.05	0.14	0.09	0.28
SOL	0.11	0.26	0.04	0.42
Other Accounting Information				
CF	0.54	0.70	0.18	1.42
NSA	0.04	0.37	0.01	0.41
VM	2.14	0.43	0.59	3.17

This table shows the mean of the extent of use of accounting information across different time horizons. FPER – financial performance, FPOS – financial position, OACC – other accounting information, ACC – accounting information. ACT – activity ratios, EARN – earnings, EXP – expenses, OFPER – other financial performance, PRATIO – profitability ratios, REV – revenue, ASS – asset, EQ – equity, LIAB – liability, LIQ – liquidity ratios, OFPOS – other financial position, SOL – solvency ratios, CF – cash flow, NSA – non-specific accounting information, VM – Valuation models.

The results in Table 15 reveal that references to accounting information are mostly historical with a mean value of 20.28% compared to a mean of 15.03% for forward-looking references. The mean percentage of references to historical accounting information is higher for all categories of accounting information with the exception of earnings and valuation models. Of the total mean references of 8.67% for the earnings category, 4.33% are forward-looking references while 4.04% are historical references. The higher mean value for the latter sub-component may reflect the importance of earnings for valuation purposes. This is further reinforced by the higher percentage of references to forward-looking information for the valuation model category with a mean percentage of 2.14% compared to 0.43% for historical references and 0.59% for non-time specific references.

6.2.3 Discussion of the descriptive statistics on the use of accounting information

The summary statistics presented above reveal significant use of accounting information by analysts. These values are in line with previous studies such as the 35% reported in Abdolmohammadi *et al*. (2006) for the category of financial information. Similarly Breton and Taffler (2001) find that on average profitability accounts for 35% of the references in analysts' reports, while financial position only accounts for 8.5%, making a total of about 43.5%. Rogers and Grant (1997) and Fogarty and Rogers (2005) also found that financial and operating information are most frequently occurring type of information in analysts' reports as they account for 51% of the information units. In those studies, the coding scheme is significantly different from this present study. For instance, while this thesis focuses only on accounting information, previous studies have examined both accounting and non-accounting information, with the former often included in a *financial* or *operating* information category that contains other information beyond accounting. This may explain the higher percentages reported in these studies.

Unlike the previous studies cited above, the sampled reports used for this study are results reports which are issued following the release of the financial statements. Although it may be expected that such reports would contain a higher proportion of accounting information than previous research have documented, but the results do not support this. The mean values below 50% imply that the majority of the content of analysts' results reports relates to non-accounting information. Prior studies indicate that analysts refer to a variety of non-financial information including information about firms' corporate governance, critical success factors, intellectual capital, management, risks, strategy, strengths, weaknesses and opportunities (see for instance Breton and Taffler 2001; Abdolmohammadi *et al*. 2006; Flöstrand 2006; Flöstrand and Ström 2006; García-Meca and Martínez 2007; Nielsen 2008; Abhayawansa and Guthrie 2012).

The results also suggest that, within accounting information, financial performance information is most relevant to analysts. Again, this is consistent with previous studies of the content of analysts' reports. Rogers and Grant (1997) found that EPS is the single most cited information in analysts' reports. Relatedly,

Breton and Taffler (2001) report that profitability is the dominant theme in analysts' reports. Similarly, Previts *et. al.* (1994), report that income statement related information dominates analysts' reports, particularly references to earnings and revenue. They argue that this finding supports the link between earnings and market values in the 'value relevance' literature.

Interestingly, the result presented above reveals that earnings are not the sole financial performance information item referred to by analysts. Revenue is equally important and, to a lesser extent, profitability ratios such as profit margins. The higher mean value for revenue of 11% compared with the mean of 8.67% for earnings suggests that revenue information is referred to most often than earnings, despite not featuring in most value relevance studies or generally having a lower value relevance than earnings (Barton *et al*, 2010). This is similar to findings in Asquith *et al.* (2005), which reported about 41.7% remarks on revenue compared to 35.8% for earnings in their sample of over 1126 reports. A closer examination of analysts' reports reveals that frequent references to revenues is not associated with greater use of related valuation models such as *price/sales ratios,* as earnings-based valuation models remained the prevailing models. Thus, the relevance of revenue to analysts does not derive from its use as a direct input to valuation models, which is usually the perspective of value relevance studies. Barker *et al.* (2012) argue that value relevance studies have a narrow definition of *relevance* which doesn't fully describe usefulness of information to investors such as fund managers. Given the view that revenues are incomplete financial performance measures (Suozzo, *et al.* 2001) and the fact that they are not used as direct inputs to analysts' valuation models, why are they referred to more often than earnings? Bricker *et al.* (1995) and Barker and Imam (2008) provide possible explanations. Both studies examined the concept of earnings quality and found that analysts often used revenues to assess earnings quality. Barker and Imam (2008, p.319) revealed that analysts' assess sources of revenue in relation to earnings and usually perceive revenues from core operations to be associated with higher earnings quality. Similarly, Bricker *et al.* (1995) suggest that earnings quality is associated with predictability of revenue. Consistent with Barker *et al.* (2012), relevance of information for analysts goes beyond the value relevance perspective and involves the usefulness of such information (such as revenues) in assessing/predicting earnings.

The results for the financial position and other accounting information items such as the cash flows category is also in line with other studies which document fewer references to balance sheet information. Breton and Taffler (2001) found only 8.5% references to financial position-related information; this is much lower in Rogers and Grant at only 3% with 1% references to cash flow. Similarly, Previts *et al*. (1994) suggest that most financial position-related references relate to *assets*, *capital* and *debt*, which is consistent with the findings presented above. An earlier study by Govindarajan (1980) also shows that analysts place much emphasis on earnings than cash flows in their reports. This is also reflected in the widespread use of earnings based (rather than cash flow based) valuation models in analysts' reports (Bradshaw 2002, Demirakos *et al*. 2004).

With regard to the time orientation of the statements discussing accounting information, the overall results show that references to accounting information is both for valuation and stewardship purposes. More frequent forward-looking references to earnings and valuation model categories suggest that analysts' focus more on the future values of accounting information that is predominantly used as direct inputs into valuation models (such as earnings). This suggests that forward-looking information is most useful for equity valuation. Further examination of analysts' reports reveals that discussion of historical accounting information often occurs in the context of assessing managerial performance. Actual results are often compared with historical trends, prior managerial estimates and analysts' consensus estimates. Most results reports begin with analysis of historical performance, including statements such as:

"Revenue of $2,495M was **above our estimate** of $2,442M and **above consensus** of $2,488M."

(Source: Canaccord report on EBay, January, 2011)

"MYL reported adjusted diluted 4Q10 EPS of $0.45 which **was in-line with consensus** and $0.01m **higher than our estimate**."

(Source: Collins Stewart report on Mylan Inc, February, 2011)

These quotes reveal that information disclosed in analysts' reports is not necessarily used as direct input into valuation models but is also useful for assessing managerial performance by comparing actual results to previous consensus, analysts' or managerial estimates. Prior research suggest that analysts'

judgement of the quality of accounting information and the extent to which they rely on this, is based in part on the quality of management (Holland 2004, Barker and Imam 2008).

6.3 Univariate analysis

6.3.1 Pearson correlation coefficient

Table 16 presents the Pearson correlation coefficients for the independent and the dependent variables. Analysis henceforth excludes

Table 16: Correlation matrix

	ΔP_EPS	Size	Leverage	Risk	Growth	Intan	MTB	ACC	FPER	FPOS	EARN	PRATIO	REV	EQ	CF
ΔP_EPS	1.00														
Size	-0.03	1.00													
Leverage	-0.02	-0.15**	1.00												
Risk	0.06	-0.13**	-0.17***	1.00											
Growth	-0.09	0.22***	-0.22***	-0.06	1.00										
Intan	-0.15**	0.02	0.13**	-0.23***	0.12**	1.00									
MTB	-0.05	-0.06	0.26***	0.01	0.06	-0.04	1.00								
ACC	0.08	-0.11*	0.19***	-0.05	-0.17***	0.05	-0.05	1.00							
FPER	0.09	-0.04	0.12*	-0.07	-0.13*	0.07	-0.06	0.92**	1.00						
FPOS	0.09	-0.13**	0.16***	0.02	-0.15**	-0.11*	0.04	0.23***	-0.04	1.00					
EARN	0.02	-0.03	0.23***	-0.10	-0.25***	0.06	-0.09	0.67***	0.64***	0.07	1.00				
PRATIO	0.17***	-0.11*	-0.08	0.00	-0.16***	-0.03	-0.02	0.41***	0.47***	-0.09	0.06	1.00			
REV	0.02	0.06	-0.02	-0.02	0.12**	0.08	-0.04	0.62***	0.73***	-0.11	0.15**	0.17***	1.00		
EQ	0.08	-0.02	0.18***	-0.16***	-0.04	0.01	-0.01	0.17***	-0.03	0.74***	0.1*	-0.13**	-0.08	1.00	
CF	0.02	-0.07	0.13**	0.05	-0.13**	0.10	-0.02	0.31***	0.14**	0.3***	0.11*	0.01	0.07	0.16***	1.00

This tables shows the Pearson correlation coefficients of the independent variables and the first level categories of accounting information. ΔP_EPS is the percentage change in EPS from prior year, Size is the natural logarithm of market value for the fiscal year end 2010, Leverage is the long term debt to total asset ratio. Risk is the standard deviation of EBIT over five years scaled by average assets over the time period, Growth is the cumulative annual growth rate in sales per share over the previous five years from 2006-2010. Intan is the proportion of intangible assets over total assets. MTB is the ratio of market value of equity to book value of equity for the fiscal year end 2010, ACC – accounting information, FPER – financial performance, FPOS – financial position, EARN – earnings, PRATIO – profitability ratios, REV – revenue, EQ – equity, CF – cash flow. * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level.

According to the results presented in Table 16, several independent variables are significantly correlated. Particularly, the results reveal that percentage change in EPS from prior year is negatively associated with proportion of intangibles in the balance sheet. This suggests that firms with increases in EPS have less intangibles. Leverage is negatively correlated with risk (-0.17 at 1% level of significance) and firm size (-0.15 at 5% level) and positively related with intangibles (0.13 at 5% level) and market to book ratio (0.26 at 1% level). This implies that highly levered firms are smaller, have less volatile earnings, more intangibles and higher market to book ratios. Moreover, risk is negatively correlated with intangibles and size with correlation coefficients of -0.23 and -0.13 respectively and significant at the 1% and 5% levels. These associations are useful in understanding the profile of the sampled firms. However, the figures are all below 0.5, suggesting that the associations are not strong enough to pose a multi-collinearity threat in the multivariate analysis. Kennedy (2008, p. 196) suggest that a value of 0.8 and above is useful indication of collinearity between variables.

Table 16 also shows that firm attributes are correlated with analysts' use of accounting information. At the high level categories, accounting information is positively associated with leverage and negatively associated with sales growth and firm size at the 1% significant levels, indicating that analysts' reports contain more accounting information for firms with higher leverage, smaller size and lower sales growth. At the lower level, the positive association between extent of use of accounting information and leverage is maintained for the financial position and performance categories and sub-categories such as earnings, equity and cash flow. The association between size and extent of use of financial performance categories are largely insignificant with the exception of the profitability ratios category, which is negatively correlated with firm size at the 10% level, while size is negatively correlated with extent of use of financial position related information with a coefficient of -0.13, which is significant at 5% level. This suggests that firm size may not be an influential factor in explaining the extent of variation in the use of financial performance-related information in analysts' reports but for financial position related information. Finally, as with the higher level, references to financial performance related items, earnings, profitability ratios and cash flows are negatively correlated with growth at the 10%, 1% and 5% levels respectively, but positively associated with extent of use of revenue at the 5%

level. This suggests that analysts following companies with high growth in sales produce reports which contain more references to revenue.

Level of percentage change in EPS from previous year is not significantly correlated with extent of use of accounting information or its sub-components, with the exception of the *PRATIO* category. Extent of discussion of profitability ratios in the reports is positively associated with changes in EPS at the 1% level (providing support for H_1).

The correlation coefficient between higher and lower level categories of the dependent variables are large and significant (see for instance, ACC and FPER). However, this is due to the fact that higher level categories are combinations of lower level categories and as such, high correlations are expected. Interestingly, there are some significant associations among the lower level variables. For instance, there is a significant positive association between extent of discussion of earnings (EARN) and profitability ratios (PRATIO) and extent of discussion of revenue (REV), suggesting that analysts discussion of earnings or profitability ratios is usually accompanied with discussion of revenue. This supports the view that revenue is an incomplete measure of financial performance (Suozzo *et al*. 2001) and as such, analysts' do not refer to revenue in isolation but evaluate a firms' financial performance through analysis of different accounting metrics. Such positive association is also observed for the *EARN* and *EQ* and *CF* variables.

Table 16 examines only the association between the use of accounting information and the continuous independent variables. For categorical variables, t-test of difference in means are used to investigate the univariate relationship with accounting information. Table 17 presents the results for the industry variable.

6.3.2 Extent of use of accounting information across industries

Table 17: Use of accounting information across industries

Panel A: All references				*Panel B: Forward-looking references*				*Panel C: Historical references*			
	HTC	**LTC**	**Mean Diff**		**HTC**	**LTC**	**Mean Diff**		**HTC**	**LTC**	**Mean Diff**
Main Categories				***Main Categories***				***Main Categories***			
ACC	35.59	41.14	-5.55***	ACC	14.70	15.25	-0.55	ACC	16.96	22.45	-5.48***
FPER	26.38	31.98	-5.60***	FPER	10.58	11.54	-0.96	FPER	13.58	18.20	-4.62***
FPOS	3.85	4.34	-0.49	FPOS	1.06	1.22	-0.15	FPOS	2.05	2.56	-0.51
Sub-categories				***Sub-categories***				***Sub-categories***			
EARN	7.60	9.37	-1.76**	EARN	4.17	4.43	-0.26	EARN	3.10	4.65	-1.55***
PRATIO	4.65	6.05	-1.40**	PRATIO	1.68	2.35	-0.68**	PRATIO	2.47	3.22	-0.75*
REV	11.03	11.08	-0.06	REV	3.68	2.79	0.89**	REV	6.39	7.46	-1.07
EQ	1.72	1.50	0.22	EQ	0.60	0.52	0.09	EQ	0.89	0.78	0.11
CF	1.31	1.49	-0.18	CF	0.52	0.54	-0.02**	CF	0.56	0.79	-0.23

This table shows the mean values of the extent of use of accounting information across the two industry groups; HTC – high tech companies and LTC – low tech companies; ACC – accounting information, FPER – financial performance, FPOS – financial position, EARN – earnings, PRATIO – profitability ratios, REV – revenue, EQ – equity, CF – cash flow. * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level.

Table 17 reveals that references to accounting information partly differ across industry. Specifically, accounting information is more discussed in reports issued for low-tech companies than high-tech companies, consistent with previous studies such as Abdolmohammadi *et al*. (2006). The mean difference in references to accounting information between high-tech and low-tech companies is 5.55% and this is significant at the 1% level and supports H_6. The result is also consistent for the financial performance category and its sub categories such as earnings and profitability ratios but not for revenues.

Panels B and C present results for the forward-looking and historical references and shows an interesting pattern. Most of the differences in the use of accounting information, financial performance and earnings related information across industries are based on historical references. This supports the argument that the historical nature of traditional accounting information limits its usefulness for high-tech companies (Francis and Schipper, 1999). Although overall references to the revenue category were not significantly different across industries, forward-looking references to revenue are higher for high-tech companies and the difference is significant at the 5% level. This further supports the usefulness of forward-looking information for high-tech companies. It also demonstrates the importance of investigating different dimensions of information disclosed in narrative documents such as the analysts' reports. Beattie *et al*. (2004) acknowledged the need for analysis of accounting narratives to capture different dimensions of information disclosed. They derived a comprehensive profile which captures the time orientation, financial and quantitative/qualitative aspects of voluntary disclosures. This study examines only the time orientation dimension and the results reveal that such distinction provides additional insight into analysts' use of accounting information

6.3.3 Extent of use of accounting information across analysts

In this section, univariate tests of the relationship between analyst characteristics and the extent of use of accounting information in analysts' reports is presented and discussed. Table 18 presents results of paired sample t-tests of difference in mean references to accounting information between IB and IND-analysts. Table 19 presents the results of conditional logistic regression to test differences in the use of accounting information between IB and IND-reports. Table 20 reports independent sample t-tests of difference in the use of accounting information between CFA and Non-CFA analysts. This is followed in Table 21 by comparison of the CFA and non-CFA use of accounting information between IB and IND- analysts.

Table 18: Differences in the use of accounting information between IB and IND-analysts

Panel A: All references				*Panel B: Forward-looking references*				*Panel C: Historical references*			
	IB	**IND**	**Mean Diff**		**IB**	**IND**	**Mean Diff**		**IB**	**IND**	**Mean Diff**
Main Categories				***Main Categories***				***Main Categories***			
ACC	44.50	33.39	11.11***	ACC	19.88	10.18	9.71***	ACC	21.45	19.10	2.36
FPER	34.56	24.98	9.58***	FPER	15.16	7.16	7.99***	FPER	17.63	15.11	2.51*
FPOS	3.97	4.32	-0.36	FPOS	1.28	1.03	0.25	FPOS	2.05	2.56	-0.51
Sub-categories				***Sub-categories***				***Sub-categories***			
EARN	11.23	6.10	5.13***	EARN	6.48	2.18	4.30***	EARN	4.41	3.67	0.74
PRATIO	5.89	5.10	0.79*	PRATIO	2.36	1.81	0.55**	PRATIO	3.13	2.71	0.43
REV	12.00	10.13	1.87**	REV	4.08	2.20	1.88***	REV	7.37	6.71	0.66
EQ	1.59	1.58	0.01	EQ	0.63	0.47	0.16	EQ	0.79	0.86	-0.08
CF	1.44	1.39	0.05	CF	1.44	1.39	0.05	CF	0.73	0.67	0.06

This table reports the result of paired sample t-test of difference in mean values of the extent of use of accounting information between Investment banking analysts (IB) and analysts employed by independent research firms (IND). ACC – accounting information, FPER – financial performance, FPOS – financial position, EARN – earnings, PRATIO – profitability ratios, REV – revenue, EQ – equity, CF – cash flow. * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level.

The results in Table 18 suggest that IB and IND-analysts differ in their use of accounting information, lending support for H_7. For all the categories of accounting information, with the exception of financial position, equity and cash flow related references, the mean difference in references to categories of accounting information between IB and IND-analysts is positive and statistically significant. This implies that IB-analysts use more accounting information compared to IND-analysts. In comparing the differences across categories, it is noted that significant differences in the use of accounting information are concentrated in the *financial performance* category and its sub-components such as references to *earnings, profitability ratios* and *revenue*. Furthermore, the difference between IB and IND-analysts is concentrated in the forward-looking category as mean difference in references to historical accounting information are mostly insignificant. This implies that IB-analysts are more reliant on forward-looking information in their reports than IND-analysts. There are three possible explanations. First, IB-firms are generally larger and carry out other IB-businesses beyond equity research. Hence, analysts may have more access to superior information about the companies covered which is useful in predicting accounting information or they may have more access to managerial estimates of future accounting information. Second, given the general perception that IB-firms attract better analysts as discussed in Chapter 2 of this thesis, IB-analysts may have superior forecasting ability compared to IND-analysts, explaining why they make more references to this type of information in their reports. Third, given that IB-analysts are more incentivised to produce optimistic research output such as optimistic earnings' forecasts, forward-looking information (which is mostly estimates based on opinion) may be used strategically to support the case they want to make in terms of justifying their recommendations and target prices.

A conditional logistic regression was also performed to test differences in the use of accounting information in reports by IB and IND-analysts. The dependent variable for the conditional logistic regression model is a dichotomous variable (IB) which takes the value of 1 for reports issued by an IB-analysts and 0 for reports written by IND-analysts following the same firm. This is regressed on the measures of the extent of use of accounting information (or its sub-level categories). The use of the

conditional logistic regression provides control for biases due to sampling. Hence, the results provide a robustness check to the paired sample t-test analysis reported in Table 18. The logistic regression model estimates the probability that a report is written by an IB-analysts given the extent of use of accounting information (or its sub-components). Separate models were estimated for each category of accounting information and Table 19 presents the results.

Table 19: Differences in the use of accounting information between IB and IND analysts

Panel A: All references				*Panel B: Forward-looking references*				*Panel C: Historical references*			
	Coeff.	**Odds Ratio**	**Significance**		**Coeff.**	**Odds Ratio**	**Significance**		**Coeff.**	**Odds Ratio**	**Significance**
Main Categories				***Main Categories***				***Main Categories***			
ACC	0.06	1.06	0.00	ACC	0.24	1.27	0.00	ACC	0.01	1.01	0.13
FPER	0.06	1.06	0.00	FPER	0.32	1.37	0.00	FPER	0.02	1.02	0.08
FPOS	-0.03	0.97	0.37	FPOS	0.12	1.13	0.15	FPOS	-0.05	0.96	0.34
Sub-categories				***Sub-categories***				***Sub-categories***			
EARN	0.19	1.21	0.00	EARN	0.63	1.88	0.00	EARN	0.05	1.05	0.12
PRATIO	0.06	1.06	0.08	PRATIO	0.14	1.15	0.03	PRATIO	0.06	1.06	0.21
REV	0.03	1.03	0.05	REV	0.31	1.37	0.00	REV	0.01	1.01	0.42
EQ	0.00	1.00	0.98	EQ	0.17	1.18	0.20	EQ	-0.05	0.95	0.61
CF	0.01	1.01	0.83	CF	0.11	1.12	0.36	CF	0.04	1.04	0.69

The dependent variable is IB – a dummy variable which takes the value of 1 for IB-reports and 0 for IND-reports. The independent variables are the extent of use of accounting information and its subcomponents. The table shows the coefficient, odds ratio and their statistical significance. ACC – accounting information, FPER – financial performance, FPOS – financial position, EARN – earnings, PRATIO – profitability ratios, REV – revenue, EQ – equity, CF – cash flow.

The *coefficient* measures changes in the log of odds that the report is issued by an IB-analyst for a percentage increase in the use of accounting information (or its sub-categories). For a more intuitive interpretation, the *odds ratio* is also reported and shows the increase in the odds of the report being issued by an IB-analysts given a percentage increase in the use of accounting information. The significance level, based on *Wald* test is also reported.

With the exception of the financial position and equity category, the coefficients are all positive and the odds ratios greater than one. This indicates that increases in the use of accounting information, increases the likelihood that a report is issued by an IB-analysts. As in the paired sample t-test, the results suggest that IB-analysts' reports contain more accounting information than IND-reports. The significance column reveals that the differences between IB and IND-reports are statistically significant for the accounting, financial performance, earnings, profitability ratios and revenue categories and the differences are mostly related to forward-looking references. The results in Table 18 are supported by the results from the conditional logistic analysis which adjusts for the bias introduced through sampling.

Table 20: Differences in the use of accounting information between CFA and non-CFA qualified analysts

Panel A: All references				*Panel B: Forward-looking references*				*Panel C: Historical references*			
	CFA	**Non-CFA**	**Mean Diff**		**CFA**	**Non-CFA**	**Mean Diff**		**CFA**	**Non-CFA**	**Mean Diff**
Main Categories				***Main Categories***				***Main Categories***			
ACC	39.34	39.57	-0.23*	ACC	16.46	14.58	1.88	ACC	19.35	21.46	-2.11
FPER	30.04	30.26	-0.21	FPER	12.29	10.76	1.54*	FPER	15.54	17.42	-1.88
FPOS	4.41	3.93	0.48	FPOS	1.43	0.98	0.45**	FPOS	2.35	2.39	-0.04
Sub-categories				***Sub-categories***				***Sub-categories***			
EARN	8.86	8.84	0.02	EARN	4.68	4.26	0.43	EARN	3.84	4.31	-0.47
PRATIO	5.77	5.39	0.38	PRATIO	2.37	1.95	0.42	PRATIO	2.90	2.99	-0.09
REV	10.14	11.97	-1.82*	REV	3.18	3.25	-0.06	REV	6.19	7.86	-1.67*
EQ	1.56	1.62	-0.07	EQ	0.55	0.58	-0.03	EQ	0.82	0.82	0.00
CF	1.59	1.34	0.25	CF	0.71	0.42	0.29**	CF	0.67	0.75	-0.09

This table reports test of differences in the mean extent of use of accounting information between CFA and Non-CFA analysts. ACC – accounting information, FPER – financial performance, FPOS – financial position, EARN – earnings, PRATIO – profitability ratios, REV – revenue, EQ – equity, CF – cash flow. * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level

The results in Table 20 reveal only minor differences in the use of accounting information between chartered and non-chartered financial analysts. Non-chartered financial analysts make more references to accounting information and revenue (particularly historical references) but this is only marginally significant at the 10% level. In splitting the references between forward-looking and historical, the differences are mostly positive for the forward looking references (with only the financial performance, financial position and cash flow category significant at the 10% and 5% level) and negative for historical references (although mostly insignificant). This suggests that CFA qualified analysts make more references to forward-looking accounting information than their non-CFA colleagues. This evidence may reflect an increased knowledge of the limitations of historical accounting information or increased ability to forecast accounting information by CFA-analysts. De Franco and Zhou (2009) examined the difference in performance between CFA and non-CFA charter holders and suggest that charter holders acquire more financial knowledge and skill, leading to greater performance. The present study suggests that this additional financial knowledge and skills acquired through the CFA program may also influence analysts' use of accounting information. Specifically, CFA-analysts are either more adept at forecasting or are more aware of the limitations of historical accounting information and thus rely more on forward-looking accounting information compared to their non-CFA counterpart.

The difference between CFA and non-CFA are cross-tabulated for IB and IND-analysts separately and the results are presented in Table 21

Table 21: Difference in the use of accounting information between CFA and non-CFA analysts and IB vs IND-analysts

Panel A: IB analysts

All references				*Forward-looking references*				*Historical references*			
	CFA	Non-CFA	Mean Diff		CFA	Non-CFA	Mean Diff		CFA	Non-CFA	Mean Diff
Main Categories				*Main Categories*				*Main Categories*			
ACC	45.73	43.33	2.41	ACC	19.82	19.94	-0.12	ACC	22.94	20.05	2.88
FPER	35.96	33.23	2.73	FPER	14.96	15.34	-0.38	FPER	19.16	16.18	2.98
FPOS	3.98	3.95	0.03	FPOS	1.34	1.23	0.11	FPOS	2.15	2.29	-0.13
Sub-categories				*Sub-categories*				*Sub-categories*			
EARN	11.77	10.72	1.06	EARN	6.53	6.43	0.09	EARN	4.84	4.00	0.84
PRATIO	6.27	5.53	0.74	PRATIO	2.40	2.32	0.08	PRATIO	3.45	2.84	0.61
REV	11.88	12.11	-0.23	REV	3.49	4.65	-1.16*	REV	7.81	6.95	0.86
EQ	1.51	1.66	-0.15	EQ	0.51	0.75	-0.24	EQ	0.82	0.76	0.06
CF	1.38	1.50	-0.12	CF	0.61	0.57	0.04	CF	0.66	0.79	-0.14

Panel B: IND analysts

All references				*Forward-looking references*				*Historical references*			
	CFA	**Non-CFA**	**Mean Diff**		**CFA**	**Non-CFA**	**Mean Diff**		**CFA**	**Non-CFA**	**Mean Diff**
Main Categories				***Main Categories***				***Main Categories***			
ACC	31.20	36.09	-4.88*	ACC	12.18	9.62	2.56**	ACC	14.79	22.77	-7.98***
FPER	22.51	27.51	-5.00**	FPER	8.89	6.51	2.38***	FPER	10.94	18.56	-7.62***
FPOS	4.96	3.91	1.05*	FPOS	1.55	0.75	0.80***	FPOS	2.61	2.48	0.13
Sub-categories				***Sub-categories***				***Sub-categories***			
EARN	5.16	7.11	-1.95*	EARN	2.33	2.24	0.09	EARN	2.58	4.60	-2.02**
PRATIO	5.13	5.26	-0.12	PRATIO	2.32	1.61	0.72**	PRATIO	2.21	3.14	-0.93*
REV	7.94	11.83	-3.90***	REV	2.80	1.96	0.84**	REV	4.14	8.70	-4.56***
EQ	1.62	1.59	0.03	EQ	0.61	0.43	0.18	EQ	0.81	0.87	-0.06
CF	1.85	1.19	0.66*	CF	0.84	0.29	0.55***	CF	0.68	0.72	-0.04

This table shows the mean values of the extent of use of accounting information across CFA and non CFA analysts after splitting the overall sample into two (IB vs IND-analysts). Panel A shows the results for the IB sub-sample and Panel B for the IND sub-sample. . ACC – accounting information, FPER – financial performance, FPOS – financial position, EARN – earnings, PRATIO – profitability ratios, REV – revenue, EQ – equity, CF – cash flow. * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level

Table 21 show that, for the IB-analysts' sub-sample, there is no significant difference in the use of accounting information between CFA and non-CFA analysts. However, for the IND-analysts sub-sample, there is a marked difference in the use of accounting information between CFA and non-CFA analysts. Hence, H_8 is only supported for the IND-analysts' sub-sample. For this sub-sample, non-CFA analysts make more references to accounting information in general and to financial performance related information in their reports compared to CFA-analysts as the overall mean difference between the CFA and non-CFA analysts is negative for this category and its sub-components. However, CFA-analysts use more financial position and cash flow related information compared to non-CFA analysts.

Comparing the results across forward-looking and historical references, the results show that the mean difference in the use of accounting information for the forward-looking sub-sample is positive across the categories of accounting information and significant in most cases. This indicates that CFA-analysts make more references to forward-looking accounting information while non-CFA analysts make more references to historical accounting information as revealed by the negative mean differences in the final column. This pattern supports the findings previously reported for the overall sample. Furthermore, the difference between the IB and IND sub-samples suggests that the difference in the use of accounting information between CFA and non-CFA holders is greater for analysts' employed by IND analyst-firms than analysts' employed by IB analysts-firms[34]. There are at least two reasons why this may occur. First, given the larger size and resources of IB-firms[35], IB-analysts may benefit from improved in-house training such that the gap in financial knowledge and skill is compensated for through in-house training. Second, investment banks and brokerage houses are able to recruit more experienced, qualified analysts, given their resources, size, reputation and influence. Hence, they are more likely to employ analysts with more financial knowledge obtained through other means apart from the CFA program such as prior experience, chartered

[34] Other factors that may influence the difference in use of accounting information across the IB and IND sub-sample are controlled for as a matched sample is used.

[35] See discussion of differences between IB and IND analysts in Chapter 2.

accountancy qualifications, PhD or MBA qualification and thus the difference in financial knowledge and skills, which may explain usage of accounting information, is minimised for this group.

6.4 Multivariate test of factors which explain the extent of use of accounting information

This section presents and discusses the results of multivariate analysis. Tobit regression is used to estimate the association between the extent of use of accounting information (and sub-categories) and the independent variables. The extent of use of accounting information (measured using content analysis) is the dependent variable and the independent variables are the company-, analyst- and report-specific characteristics previously discussed. Table 22 presents the results for the first level categories (i.e. extent of use of accounting information, financial performance and financial position related information), while Table 23 and Table 24 presents result for the lower level categories (i.e. extent of use of earnings, revenue, profitability ratios, equity and cash flow related information).

Table 22: Tobit regression results of accounting information on company, analysts and report characteristics I

	Dependent Variable								
	ACC	FL_ACC	HIS_ACC	FPER	FL_FPER	HIS_FPER	FPOS	FL_FPOS	HIS_FPOS
Intercept	33.47***	17.17***	7.42	18.73*	8.07**	5.3	6.08**	0.57	1.03
Company									
P_EPS	0.70*	0.14	0.55	0.72**	0.20	0.53	0.12	0.02	0.07
Size	-0.27	-0.80	0.98	0.44	-0.26	0.96	-0.36	-0.17	-0.05
Leverage	18.65**	8.83*	8.71	12.01	4.18	6.73	2.61	3.66**	-0.30
Risk	1.72	-7.50	23.42	-3.38	-5.74	13.66	-0.43	0.85	3.31
Growth	-16.39*	-4.74	-14.32	-8.80	-1.62	-10.01	-5.20	-2.53	-3.52
Intan	10.68**	-1.30	13.39***	12.03**	0.98	12.17***	-2.03	-1.42	-0.05
MTB	0.45	0.40**	-0.02	0.46	0.41***	-0.03	-0.09	0.00	0.01
LTC	3.73*	-2.29**	6.58***	5.41***	-0.52	5.84***	-0.29	-0.33	0.15
Analyst									
CFA	1.56	2.04*	-0.07	1.26	1.16	0.34	0.29	0.27	-0.15
IB	6.75***	8.39***	-1.04	5.53***	7.15***	-0.83	-0.16	0.41	-0.40
Report									
POS	2.17	0.72	1.26	0.78	-0.33	1.1	0.32	-0.29	0.41
WC	-0.004***	0.000	-0.003***	-0.004***	0.000	-0.003***	0.001***	0.001***	0.001**
F	8.76***	9.35***	4.42***	9.32***	9.39***	4.83***	2.16**	3.43***	1.75*
Pseudo R2	2.93%	4.76%	1.80%	2.94%	4.87%	1.95%	2.08%	4.44%	1.63%
N	247	2.47	2.47	2.47	2.47	2.47	247	251	251

This table presents the results of a Tobit regression of extent of use of accounting information on companies, analysts and reports characteristics. The suffix _FL and _HIST represents categories of forward-looking and historical references to accounting information. ACC – accounting information, FPER – financial performance, FPOS – financial position. These variables are the derived from the content analysis process described in chapter 5. ΔP_EPS is the percentage change in EPS from prior year, Size is the natural logarithm of market value for the fiscal year end 2010, Leverage is the long term debt to total asset ratio, Risk is the standard deviation of EBIT over five years scaled by average assets over the time period, Growth is the cumulative annual growth rate in sales per share over the previous five years from 2006-2010, Intan is the proportion of intangible assets over total assets. MTB is the ratio of market value of equity to book value of equity for the fiscal year end 2010, LTC is a dummy variable which takes the value of 1 for firms in low tech industry and 0 for firms in high tech industry. CFA is dummy variable which takes the value of 1 for reports by at least one CFA qualified analysts and 0 for reports authored by non-CFA analysts. IB is dummy variable which takes the value of 1 for reports by IB-analysts and 0 for IND-analysts. POS is dummy variable which takes the value of 1 for reports with a positive recommendation and 0 otherwise. WC is the word count for the reports excluding regulatory disclosures, tables and figures. ***/**/* means significance at 1%, 5% and 10% respectively.

Table 23: Tobit regression results of accounting information on company, analysts and report characteristics II

	Dependent Variable								
	EARN	FL_EARN	HIS_EARN	REV	FL_REV	HIS_REV	PRATIO	**FL_PRATIO**	**HIS_PRATIO**
Intercept	1.77	-0.64	0.43	5.04	1.85	-0.45	8.53**	4.08*	2.55
Company									
P_EPS	0.02	0.05	0.004	0.22	0.04	0.17	0.43**	0.17	0.31**
Size	0.50	0.20	0.36	0.44	-0.01	0.58	-0.43	-0.34*	-0.01
Leverage	11.01***	9.03***	2.06	0.72	-3.89**	2.86	-4.26	-3.96**	-2.67
Risk	-7.24	-6.69	0.61	14.64	-1.83	21.21	-9.14	0.86	-5.67
Growth	-11.03***	-1.23	-9.78***	11.18	1.20	7.95	-8.27**	-2.13	-4.81
Intan	2.71	0.08	2.98	6.39*	0.59	6.23**	1.30	-0.41	2.43
MTB	-0.19	0.03	-0.22**	0.45**	0.30***	0.18	0.30**	0.02**	0.00
LTC	1.35	-0.22	1.90***	1.26	-1.30**	2.61**	0.90	0.47	0.96
Analyst									
CFA	1.58**	0.99**	0.51	-1.42	-0.38	-0.87	0.05	0.17	0.10
IB	3.44***	3.91***	0.13	0.15	2.10***	-1.27	0.6	0.95**	-0.02
Report									
POS	1.55**	0.55	1.12**	-0.45	-0.7	0.13	-0.96	-0.79*	-0.35
WC	-0.002***	-0.001***	-0.001***	-0.001***	0.001***	-0.001***	0.000	0.001***	0.000
F	15.02***	11.63***	5.5***	3.09***	3.98***	3.36***	2.47***	3.01***	1.42
Pseudo R2	6.69%	8.70%	3.89%	1.36%	3.83%	1.42%	1.63%	3.25%	1.21%
N	247	247	247	247	247	247	247	251	251

This table presents the results of a Tobit regression of extent of use of accounting information on companies, analysts and reports characteristics. The suffix _FL and _HIST represents categories of forward-looking and historical references to accounting information. EARN - earnings, PRATIO – profitability ratios, REV – Revenues. These variables are the derived from the content analysis process described in chapter 5. ΔP_EPS is the percentage change in EPS from prior year, Size is the natural logarithm of market value for the fiscal year end 2010, Leverage is the long term debt to total asset ratio, Risk is the standard deviation of EBIT over five years scaled by average assets over the time period, Growth is the cumulative annual growth rate in sales per share over the previous five years from 2006-2010, Intan is the proportion of intangible assets over total assets. MTB is the ratio of market value of equity to book value of equity for the fiscal year end 2010, LTC is a dummy variable which takes the value of 1 for firms in low tech industry and 0 for firms in high tech industry. CFA is dummy variable which takes the value of 1 for reports by at least one CFA qualified analysts and 0 for reports authored by non-CFA analysts. IB is dummy variable which takes the value of 1 for reports by IB-analysts and 0 for IND-analysts. POS is dummy variable which takes the value of 1 for reports with a positive recommendation and 0 otherwise. WC is the word count for the reports excluding regulatory disclosures, tables and figures. ***/**/* means significance at 1%, 5% and 10% respectively.

Table 24: Tobit regression results of accounting information on company, analysts and report characteristics III

	Dependent Variable					
	EQ	FL_EQ	HIS_EQ	CF	FL_CF	HIS_CF
Intercept	0.73	-0.3	-1.67	-0.81	0.22	-3.60
Company						
P_EPS	0.13	0.06	0.08	0.06	0.05	-0.01
Size	-0.15	-0.15	-0.09	-0.11	-0.30	0.13
Leverage	3.70	3.17	2.27	2.67	1.79	0.83
Risk	-22.32***	-10.78	-15.61**	8.04	-10.33*	15.54**
Growth	-3.14	-3.04	-0.71	-3.83	-1.62	-4.56*
Intan	-0.63	-1.36	1.69	2.80**	2.52**	1.50
MTB	0.04	0.05	0.03	-0.05	-0.03**	0.00
LTC	-0.67	-0.55	-0.29	-0.01	-0.59	0.39
Analyst						
CFA	-0.01	-0.19	-0.1	0.49	0.71*	0.11
IB	0.17	0.35	0.15	-0.05	0.14	-0.05
Report						
POS	0.39	0.06	0.63	0.24	0.45	-0.31
WC	0.001***	0.001***	0.001***	0.001**	0.001***	0.000
F	1.99**	1.75*	2.36***	2.68***	2.99***	1.21
Pseudo R2	2.94%	4.20%	4.11%	3.11%	7.21%	2.73%
N	247	247	247	247	251	251

This table presents the results of a Tobit regression of extent of use of accounting information on companies, analysts and reports characteristics. The suffix _FL and _HIST represents categories of forward-looking and historical references to accounting information. EQ - Equities and CF – cash flows. These variables are the derived from the content analysis process described in chapter 5. ΔP_EPS is the percentage change in EPS from prior year, Size is the natural logarithm of market value for the fiscal year end 2010, Leverage is the long term debt to total asset ratio, Risk is the standard deviation of EBIT over five years scaled by average assets over the time period, Growth is the cumulative annual growth rate in sales per share over the previous five years from 2006-2010, Intan is the proportion of intangible assets over total assets. MTB is the ratio of market value of equity to book value of equity for the fiscal year end 2010, LTC is a dummy variable which takes the value of 1 for firms in low tech industry and 0 for firms in high tech industry. CFA is dummy variable which takes the value of 1 for reports by at least one CFA qualified analysts and 0 for reports authored by non-CFA analysts. IB is dummy variable which takes the value of 1 for reports by IB-analysts and 0 for IND-analysts. POS is dummy variable which takes the value of 1 for reports with a positive recommendation and 0 otherwise. WC is the word count for the reports excluding regulatory disclosures, tables and figures. ***/**/* means significance at 1%, 5% and 10% respectively.

Overall, the regression models are statistically significant at the 1% level in most cases. Thus, the null hypotheses that the coefficients of the independent variables are jointly equal to zero can be rejected. In other words, variations in the independent variables can jointly explain the variations in the extent of use of accounting information (and its sub-categories) in analysts' reports. The pseudo R^2 values also range from 1.21% (for the *HIS_Pratio* model) to 8.70% (for the *FL_EARN* model). The significance of the individual variable coefficients are discussed separately below.

6.4.1 Company characteristics

Financial Performance

Percentage change in EPS is used to measure financial performance in the regression models. The coefficient on percentage change in EPS from prior year is statistically significant in only four models, i.e. *ACC*, *FPER*, *PRATIO* and *PRATIO_HIS*. The coefficients are positive and significant at the 10% level or better and support H_1. This implies that percentage increase in EPS is associated with greater use of the financial performance-related information. Further analysis splitting into lower level categories and time orientation, indicates that the positive result for the financial performance category is driven principally by greater references to past-oriented profitability ratios such as margins and accounting returns. This implies that firms with a positive trend in EPS attract more comments on their financial performance by analysts and supports the argument from the value relevance literature that positive earnings are more persistent and more value relevant than negative earnings. Moreover, it shows a tendency by analysts' to highlight positive financial performance in the reports.

Risk

The conjecture is that higher volatility in operating earnings signifies greater operating risk and implies less predictable earnings. Consequently, a negative coefficient for the risk variable is generally expected for references to earnings related information but the effect on other types of accounting information is not pre-stated. The results reveal that volatility of operating earnings is not significantly associated with references to accounting information or financial performance

related information at the higher and lower levels. However, the coefficient is negative and significant for the *EQ* and *CF* models and their time-dimension, suggesting that analysts' reports for companies with higher risk levels contain less references to equity and cash flow information. Hence, H_2 is only partially supported by the data.

Firm Size

The univariate results in Table 16 suggest that firm size (proxied by the natural log of market value) is negatively associated with the extent of use of accounting information at the broad level. However, this is not maintained in a multivariate setting as the coefficient on firm size is insignificant across models with the exception of FL_PRATIO, though this is only marginal. Thus, there is no support for H_3. This may be partly due to the fact that the sampled companies, being part of the S&P 500 index, are relatively large companies and thus the information environment is not significantly different across them to result in significant differences in accounting information.

Growth

Hypothesis H_4 tests whether there is an association between sales growth (Growth), market-book (MTB) ratios and the extent of use of accounting information. The hypothesis is supported for the *Growth* and *MTB* variables. With regard to the sales growth variable, the association between the extent of use of accounting information for the broad level categories (e.g. financial performance and financial position) are not significant, although it is negative and significant at the 10% level for the *ACC* model. However, at the lower level categories, the coefficient on growth is negative and highly significant for the EARN and HIS_EARN models and PRATIO and HIS_CF models. To the extent that market-book ratios proxy for growth opportunities as previously argued in extant literature, the results provide additional support for hypotheses H_4. The coefficient for the market-book ratio is positive and significant for models relating to forward looking accounting information (e.g *FL_ACC, FL_FPER, FL_REV and FL_PRATIO*) and it is negative for HIS_EARN model with a coefficient of -0.22, which is significant at the 5% level. Precisely, for firms with higher market-book ratios, analysts' reports contain more forward-looking references to accounting information, financial performance, and revenue and profitability ratios. The

conjecture is that the value for high growth firms is reflected in market values earlier than accounting data. Hence, the link between historical accounting information and market value is lesser for high growth firms. Interestingly, there is a significant positive association between extent of use of forward looking accounting information and the market to book variable while there is a negative relationship with historical references to earnings. This finding suggests that different types of accounting information may act as substitutes in certain cases and examination of lower level categories of accounting information is useful in highlighting this.

Leverage

The results in Tables 22-24 suggest that the use of accounting information is positively associated with leverage levels and the coefficient is statistically significant over several models. With the exception of the *FL_REV* and *FL_PRATIO* models, the positive association is maintained only for the forward-looking references to accounting information. This indicates that analysts use significantly more forward-looking accounting information for firms with high leverage levels. At the lower level categories, the results are mostly driven by references to earnings related information (particularly forward looking references). The choice to focus on more forward-looking information about earnings may reflect the argument previously discussed that reported earnings figures for highly levered firms may be prone to earnings management (Dechow *et al.*, 2010). Consequently, analysts make more references to forward-looking earnings information rather than reported historical earnings data. Moreover, the more frequent references to earnings information may be due in part to concerns about its quality as previous value relevance studies suggest. This study however adopts a quantitative approach to analysing the content of analysts' reports and may be limited in capturing such nuances in references.

Industry

The results reveal that the extent of use of accounting information in analysts' reports varies across industries as the coefficient on the LTC variable is significant at the 1% or 5% across several regression models, supporting H_6. At the high level categories, extent of use of accounting information and financial performance

related information is higher for companies in the low tech industry compared with those in the high tech industry. The positive coefficients on the LTC variable are maintained only for historical references (HIS_ACC has a coefficient of +6.58 a *p-value*<0.01, while HIS_FPER has a coefficient of 5.84 and a *p-value*<0.01) while forward looking references have negative coefficients (e.g. FL_ACC has a coefficient on the LTC variable of -2.29 with a *p-value*<0.05). This suggests that reports issued by analysts covering low-tech companies present and discuss more historical references to accounting information and less references to forward-looking accounting information compared to reports issued for high-tech companies. The results are also supported for lower level categories such as earnings and revenues (HIS_EARN and FL_REV and HIS_REV models). The results are supportive of the findings in Abdolmohammadi *et al*. (2006) which showed that analysts use less financial information for firms operating in intangible-intensive industries. However, this thesis provides an extension of that study as additional insight is offered by distinguishing between forward-looking and historical accounting information and the results reveal that the pattern is different based on whether the information is historical or forward looking. This distinction further provides additional support for the need to provide low level categorising of information. For further test of the effect of intangibles on the use of accounting information, the proportion of intangibles recognised in the financial statements was included as an additional variable. The extent of use of accounting and financial performance related information is significantly and positively associated with the proportion of intangibles recognised in the balance sheet. This is particularly so for historical references to accounting information and its sub-components. The higher references to historical accounting data may suggest that for companies for which intangibles are recognised in accounting statements, the gap between accounting and market value is less pronounced which enhances relevance of accounting information.

6.4.2 Analysts and report characteristics

The multivariate tests largely confirm the findings from the univariate analysis discussed in section 6.3. IB-analysts use more accounting information and particularly more forward-looking information as the coefficients for the IB variable are positive and significant mostly for the forward-looking models (FL_*ACC*, FL_*FPER*, FL_*EARN*, FL_*REV* and FL_PRATIO). The coefficient for the CFA variable are mostly insignificant except for the FL_ACC, EARN, FL_EARN and FL_CF models, for which they are positive and significant at the 10% and 5% level. This again supports the findings reported in section 6.3 that CFA-analysts use more forward-looking accounting information and this difference is pronounced for the sub-sample of IND reports. Hence, the results lend support for H_7 and H_8 and reveal the importance of distinguishing between types of analysts as overall results may mask these differences.

With regards to the recommendation types, the coefficient on POS is mostly insignificant across the regression models with the exception of the EARN and HIS_EARN models in which the coefficients are positive and significant at the 5% level. The coefficient on POS is also negative and marginally significant for the FL_PRATIO model. This suggests that use of earnings (particularly historical references) is higher in reports accompanied by a positive recommendation than reports accompanied by a negative or neutral recommendation. Breton and Taffler (2001) found that reports with a buy recommendation make more positive references to profitability. This partly supports the positive and significant coefficient observed for the EARN model. However, unlike this present study, Breton and Taffler (2001) examines the tonal words in connection to the use of accounting information and distinguishes between positive and negative references. The lack of this distinction in this study may explain why the coefficient on recommendations is not significant in most of the regression models.

6.4.3 Control variables

The coefficient for word count is statistically significant for most of the regression models. This implies that, in most cases, the use of accounting information (or its sub-categories) varies with the length of the reports. However, the magnitude of the effect on the dependent variable is very little as indicated by the very small coefficient on the *WC* variable. The results are maintained when size of reports is proxy by number of pages rather than word count.

6.4.4 Discussion of the multivariate analysis

In summary, the empirical analysis reveal that most company attributes (generally found in existing research to influence the value relevance of accounting information) were not significantly associated with analysts' use of accounting information. However, attributes that reflect the disconnection between accounting information and firm value (such as high growth firms, firms with high market-book ratio, high tech vs low tech firms) were found to influence analysts' use of accounting information. More specifically, the higher the disconnection between accounting and value (as in high market-book ratios and high growth firms), the lower the extent to which historical accounting information is relied on by analysts and the higher the references to forward-looking accounting information. Moreover, the lesser the disconnection between accounting information and value, (due to increased recognition of intangibles in the balance sheet and for sample of low tech companies), the more analysts use accounting information in their reports.

Also, it was observed that several variables were significantly associated with the extent of use of accounting information but with a different sign from that which is generally expected from the value relevance literature. For instance, the association between the extent of use of earnings in the report and leverage levels was significant and positive, which does not support prior value relevance studies that suggest lower earnings quality for firms with higher debt levels (Barth *et al*. 1998). As previously discussed, the higher number of references in this case might relate to questions about the quality of earnings rather than being indicative of its actual usage in the valuation process. Thus, a limitation of the quantitative content-

analytic approach used in this study is that it cannot reveal *how* information is used by analysts. Inferences about relevance are based only on the proportion of references to different types of accounting information. Many of the content-analytic studies of analysts' reports have used this approach. Hence, to extend the findings here and improve our understanding of the content of analysts' reports, future research may seek to use more qualitative approaches such as for instance, discourse analysis. This may provide additional insight into the content and use of information by analysts.

In summary, valuation concerns only explain part but not all the rationale for analysts' use of information. This may be due to several factors: First, value relevance studies draw upon single-person decision theory and assume rational choice, thereby ignoring behavioural aspects; behavioural theories might better explain why analysts choose to use certain information types over others. Second, value relevance studies are principally concerned with the use of accounting information for valuation, either as direct inputs into valuation models or as a summary of price-sensitive information for investors. The unstated assumption in extant content-analytic research is that relevance for valuation is reflective of use in analysts' reports. However, analysts may choose to discuss different information types in their reports for reasons beyond their use for valuation purposes. For instance, further examination of the sampled reports reveal that accounting information is used to evaluate managerial performance against previous estimates or consensus forecasts, which is more consistent with the stewardship role. Moreover, analysts' reports may also be used as a marketing tool and as such the information disclosed therein may be useful for "selling" a story about the companies covered. To the extent that information in the reports is used for purposes beyond valuation, the results may differ from those predicted by value relevance studies.

Overall, the findings in this study extend prior literature in a number of ways. First, it provides sub-level categories of accounting information and distinguishes between references to historical and forward-looking information. Beattie *et al*. (2004) advocate such detailed analysis of corporate disclosure. The results discussed above reveal that the investigation of lower level categories and two-way cross-analysis is

useful in providing a thorough analysis of variation in the use of accounting information. For instance, the examination of differences in the use of financial performance-related information revealed that revenues, rather than earnings, are the most frequently cited information item in analysts' reports. Additionally, two-way analysis revealed that the use of accounting information for high-tech firms is higher when there are forward-looking references and lower for references to historical accounting information.

This study also distinguished between IB and IND-analysts' reports and found that this distinction is important in understanding analysts' use of accounting information. Prior research has focused on how the differences between IB and IND-analysts influence the output of their research, particularly their earnings' forecasts and recommendations. Importantly, the evidence from this study suggests that IB and IND-analysts are also different in their use of accounting information. IB-analysts rely more on forward-looking information compared to IND-analysts, suggesting that they either have superior forecasting ability or have more access to superior information useful for forecasting accounting information. This implies that there may be a knowledge or resource gap between these analysts and thus the output of their analysis may be influenced by these differences, apart from their incentives to be optimistic, which has often been the focus of prior research. Furthermore, the difference in the use of accounting information also varies across analysts with or without the CFA designation. These distinctions have been largely ignored in prior literature, which focused mostly on distinguishing the output of different types of analysts, rather than understanding how their differences influence their decision making such as their use of accounting information.

6.5 Robustness tests

To assess the validity of the results, several checks were carried out. First, t-test of independence requires that the variables are normally distributed. Hence, a test of normality on the dependent variables was carried out based on the Shapiro-Wilk test of normality. This is perceived as a superior method for testing normality (Kennedy 2008). The (untabulated) results suggest that the data suffer from non-normality. However, for the independent sample t-test reported in section 6.4, the size of each of the subs-samples is greater than 30 and thus the Central Limit Theorem provides reasonable justification for relying on the results.

Second, it is possible that the results of the analysis may vary across analysts-firms as analysts use of accounting information may vary due to different in-house training, templates, styles or reporting policies. The standard regression model is based on the assumption that errors are independent and normally distributed. However, to the extent that analysts' practice with regard to accounting information is influenced by the brokerage house in which they are employed, this assumption is violated which may result in very small standard errors (Cameron and Miller, 2013). Hence, a regression model based on clustered standard errors was also estimated. This assumes that errors are uncorrelated across clusters but are correlated within clusters. The clusters in this case are the analysts-firms. The results are untabulated but in large part consistent with those reported in Tables 6.13-6.15[36]. To obtain the maximum benefit from this adjustment, cluster sizes and the number of observations within clusters should be very large (Cameron and Miller, 2013), which is not the case for the sample used. Hence, the benefits of clustering may be limited but the results show that the results are robust to alternative models. Moreover, the standard errors in the un-clustered models reported in Tables 6.13-6.15 are based on the White (1980) heteroscedastic-robust standard errors and thus controls for possible heteroskedastic concern in the data.

Third, the sample of reports used in this study is matched between IB and IND-analysts. Cram *et al*. (2009) suggest that, for choice-based and/or matched

[36] The coefficients for the independent variables in the clustered models are exactly same with those reported in Tables 6.13-6.15.

samples, the results may not be generalised for two reasons: first, choice-based samples are non-random and, second, the number of observations in the sample groups and corresponding matches are not reflective of the sizes of the corresponding groups in the population. In this study, disproportionate sampling occurs as the incidence of IND-reports is significantly lower than those of IB-reports in the population of reports available on *Investext*. To control for this, Cram *et al*. (2009) and Hosmer *et al*. (2013) advocate the use of conditional logistic regression. This estimates the likelihood conditioned on each pair of IB and IND-report. Hence, the robustness of the paired sample t-test reported in Table 18 was further assessed using conditional logistic regression and results are presented and discussed in section 6.4. As highlighted in that section, the results are robust to alternative methods.

Finally, multicollinearity checks were conducted for each of the regression models presented in Table 22-24 by computing the VIF scores. According to Wooldridge (2013), VIF scores greater than 10 indicate likely presence of multicollinearity. For each of the models, all the VIF scores were below the value of 2. Hence, multicollinearity was not a concern for the analysis.

6.6 Chapter summary

This chapter presented and discussed the results of hypothesis test which examined the factors which explain variation in the use of accounting information.

Descriptive statistics of the dependent and independent variables were presented and discussed. It was found that, consistent with previous studies, analysts' reports contain a significant number of references to accounting information. Moreover, at the lower level categories, it was documented that financial performance related information dominates and within this category, revenue information ranked higher, followed by information about earnings. Further, the descriptive statistics revealed that extent of use of accounting information does vary significantly across reports. This variation was further investigated.

Both univariate and multivariate test revealed that the use of accounting information in analysts' reports vary across some company characteristics (particularly Market-book ratios, industry classification and proportion of intangibles in the balance sheet and to a limited extent, changes in EPS and level of financial leverage), analysts' characteristics (principally whether analysts were employed by IB or IND-firms) and report characteristic (length of the report as measured by the word count).

Chapter 7: A review of the literature on analysts' bias

7.1 Chapter introduction

The second objective of this this study is to examine the factors which explain variation in the linguistic features of analysts' reports. This is addressed in Chapter 7-10 of this thesis. Chapter 7 provides a literature review of prior empirical literature on the linguistic features of analysts' reports and the literature on analysts' optimistic bias. Hypotheses are formulated in Chapter 8 while Chapter 9 provides details of the research design process, including choice of measure for the linguistic features examined. Finally Chapter 10 presents and discusses the results of the empirical analysis.

This chapter is structured as follows: section 7.2 reviews the literature on the linguistic features of analysts' reports. This is followed by a review of the literature on analysts' optimistic bias in section 7.3 which highlights the current gap and importance of examining the narrative content of analysts' reports. This is followed in section 7.4 by a discussion of the theoretical framework which explains why analysts produce optimistic research. Finally, section 7.5 concludes the chapter.

7.2 Linguistic features of analysts' reports: a review

On the premise that the work of analysts is characterised by an institutional setting in which analysts' objectivity is undermined by conflicts of interest, Fogarty and Rogers (2005) proposed and found that analysts were favourably disposed towards corporate managers and their future plans as their reports contained more positive words compared to negative words, with words related to praise and accomplishments significantly higher than words related to blame. Mokoaleli-Mokotelli *et al*. (2009) examined the linguistic features of tone, certainty and activity of analysts' reports using *Diction*. However, unlike Fogarty and Rogers (2005), the linguistic measures were used to proxy for cognitive bias such as over optimism, overconfidence and representativeness bias. The evidence suggests that reports containing new buy recommendations are more optimistic compared to new sell recommendations and such optimism is higher for reports

from analysts employed by investment banks which are affiliated with the companies.

Subsequent studies of the tone of analysts' reports have focused on the association between the tone of the reports and market variables. Kothari *et al*. (2009) examined the impact of the tone of analysts' reports on companies' cost of capital. Evidence revealed that the tone of analysts' reports is not significantly associated with cost of capital which they attribute to analysts' lack of credibility given the conflicts of interest that affect their work. However, Twedt and Rees (2012) document evidence that the tone of analysts' reports contain significant information content as reports with more positive tone are associated with higher market returns while report complexity is only associated with market returns if interacted with stock recommendations. Using a large sample of 363,952 reports, Huang *et al*. (2014) provide further evidence to support the information content of the tone of analysts' reports. The results reveal that market returns are positively associated with the tone of analysts' reports. Further analysis showed that the reaction to negative tone is significantly higher compared with positive tone, which suggests that investors are much more sensitive to the tone of negative reports, which they attribute to, investors recognition of analysts' conflict of interest which results in assigning more credibility to unfavourable reports. Using a similarly large sample of 356,463 reports, De Franco *et al*. (2015) examined the readability of analysts' reports. Their study was aimed at examining whether analysts' ability influences the readability of their reports and whether readability of the reports results in greater trading volume. The evidence indicate that readable reports were issued by high ability analysts which had more experience, issue more timely forecasts, revise forecasts more frequently, and issued more consistent forecasts and recommendations.

Thus far, these studies provide evidence that the linguistic features of analysts' reports are valuable to investors and convey information beyond the summary quantitative measures such as earnings' forecasts, recommendations and target prices. These studies also suggest that in general the tone of analysts' reports is positive, which support the evidence of prevalence of positive recommendation and forecast errors in existing literature. However, little attention has been dedicated to the study of the factors which explain variation in the tone and

readability of analysts' reports. Mokoaleli-Mokotelli *et al*. (2009) found that tone varies with recommendation types, while De Franco *et al*. (2015) found that readability vary with analysts' ability. This thesis extends this literature by examining whether the linguistic features of tone and readability of analysts' reports vary with analysts' incentive to issue optimistic research. Cue is taken from the literature on corporate reporting which have since documented the strategic use of these features by corporate managers to manage impressions about the performance of their companies. Merkl-Davies and Brennan (2007) and Brennan *et al*. (2009) listed the variety of impression management approaches documented in prior studies of accounting narratives. These include syntactical manipulation, rhetorical manipulation, attribution of organisational outcomes, thematic manipulation, selectivity, visual and presentation effects and performance comparisons.

Syntactical manipulation is concerned with a study of how the syntactic complexity of accounting reports is manipulated to obfuscate bad news (e.g. Li 2008). Rhetorical manipulation is concerned with the study of how managers make language choices or use rhetorical devices to influence impressions about their performance (e.g. Sydserff and Weetman, 2002). Attribution of organisational outcomes is concerned with the study of how managers use patterns of causal reasoning and self-serving attributions to explain corporate performance (e.g. Aerts 2005). Thematic manipulation is concerned with the study of how managers use different themes or textual features to strategically report on corporate performance (e.g. Clatworthy and Jones, 2006). Selectivity is concerned with the study of how managers strategically select performance measures to report or highlight in accounting narratives. Visual and presentation effects are concerned with the study of how managers use presentation techniques to emphasize information or manage impression of company performance (e.g. Beattie *et al*. 2008). Finally performance comparisons are concerned with the study of how managers strategically choose performance benchmarks that present their companies in a positive light. This extensive literature has provided significant empirical evidence that linguistic features are manipulated to manage impressions in corporate setting. However, despite the empirical and anecdotal evidence of analysts' optimistic bias, studies have been largely limited to study of the summary measures which accompany analysts' reports such as recommendations

and their earning's forecasts. This thesis extends the literature on impression management to the setting of analysts' reports and examines whether there is evidence of strategic reporting which is consistent with analysts' incentives to issue optimistic recommendations and forecast. Both the readability and the tone of the reports are examined.

The next section reviews the findings from existing literature on analysts' bias in order to provide relevant context for the present study. The sources of bias and related empirical evidence are reviewed.

7.3 Review of the literature on analysts' bias

Considerable academic research effort has been channelled towards investigating the issue of analysts' bias. Interest in this topic increased significantly following the dotcom crash given the public outcry over the role of analysts in the crisis that ensued. What do we know after over three decades of studies? The empirical evidence to date is somewhat conflicting as some studies found support for analysts' bias while others did not. Nevertheless, interesting facts have emerged from this literature such as the nature and sources of analysts' conflict, how this conflict influences their earnings forecasting behaviour and recommendations, mitigating factors, the influence of their employment structure and the impact of recent regulation.

The review is grouped according to the sources of analysts' conflict of interest which motivate their optimism. This ensures that the differential conflict faced by different types of analysts, which is of interest in this study, is highlighted[37]. In each subsection, the motivation for bias is discussed, empirical evidence is then reviewed, considering also the evidence from the post-regulatory setting.

37 Excellent reviews of the broad literature on analysts and the conflicts of interest they face are contained in Mehran and Stulz (2007), Ramneth *et al*. (2008), Beyer *et al*. (2010) and Bradshaw (2011).

7.3.1 Corporate managers

Early empirical evidence of the influence of corporate managers on the output of analysts' research was provided by Francis and Philbrick (1993), who documented a negative relationship between analysts' recommendations and their earnings forecasts. As recommendations take a negative direction, earnings forecasts became increasingly optimistic. In their view, the results suggest that analysts issue more positive forecasts when recommendations are negative for the purpose of pleasing management. This conclusion is based on the nature of their sample. They used forecasts issued by Value Line analysts which had associated recommendations not issued by the same analysts. Moreover, Value Line is an independent analyst-firm which does not provide investment banking or brokerage services. In this context in which recommendations are given, analysts attempt to curry favour with management through issuance of optimistic earnings forecasts to compensate for negative recommendations. Relatedly, Das *et al*. (1998) found greater forecast optimism for firms with less predictable earnings, suggesting that increased uncertainty about future earnings motivates optimistic earnings forecasts. They argue that the need for access to management's private information is greater for such companies. However, the findings from these early studies do not hold when the level of actual earnings is controlled for (Eames *et al*., 2002 and Eames and Glover 2003). Eames and Glover (2003) found no association between earnings forecast errors and earnings' predictability, contrary to Das *et al*. (1998). Moreover Eames *et al*. (2002) found that earnings forecasts are optimistic for buy recommendations and pessimistic for sell recommendations, contrary to Francis and Philbrick (1993).

Subsequent studies have proposed and documented an optimistic-pessimistic forecasting pattern by analysts who seek to maintain good relationship with corporate managers. The premise is that managers' interest in optimistic earnings' forecast is only temporary given that their ultimate intention is to ensure positive market reaction following earnings announcements. This is achieved through reporting of earnings' forecast which are initially optimistic and subsequently pessimistic (nearer to the earnings' announcement date) such that actual earnings meet or beat analysts' forecasts. Consequently, it is argued that managers engage in an "earnings guidance game" (Richardson *et al*., 2004) in which they provide analysts with earnings guidance that is beatable in order to incite positive share

price performance. The evidence thus far suggests that the optimistic-pessimistic pattern is stronger for firms whose managers have greater incentives to raise the firms' share price, such as prior to a new equity issue or exercise of stock options (Richardson *et al.*, 2004) and firms whose earnings are more difficult to predict (Ke and Yu 2006). This latter study also found that analysts who engage more in the optimistic-pessimistic earnings forecast pattern have more accurate earnings, less job turnover, are more reputable and work for larger investment banks. Whether these features are rewards from engaging with managers in the "earnings guidance game" is not conclusive as the evidence is based on statistical association and causality can only be assumed (Gassen 2014).

Feng and McVay (2010) addressed the question of analysts' bias by examining the extent to which analysts weigh managements' guidance when forecasting earnings. The study found that, prior to new equity issues and other company events requiring investment banking business such as mergers and acquisitions, IB-analysts weigh management earnings guidance more than non-IB analysts and analysts affiliated with the company weigh management guidance more than non-affiliated analysts. Moreover, they found that analysts who issue strong buy recommendations are those who overweigh management guidance, suggesting that analysts use a variety of tools to curry favour with managers.

Taken together, the empirical and anecdotal evidence suggests that corporate managers influence the output of analysts' research, given their ability to reward or punish analysts through the provision/withholding of private information about their companies (Mayew, 2008) as well as selection of their employers for investment banking businesses. However, several regulations have been enacted to limit the information advantage enjoyed by analysts who curry favour with corporate managers. As discussed in Chapter 2, Regulation Fair Disclosure (Reg FD) enacted on October 23, 2000 in the US specifically prohibited corporate managers from providing material private information about their companies to analysts without simultaneously disclosing the information publicly. However, empirical evidence on the effectiveness of this regulatory pronouncement is mixed. Some studies focused on the impact of Reg FD on analysts' information environment while others address analysts' reliance on disclosure by corporate managers. Within the first group, both Heflin *et al*. (2003) and Francis *et al*. (2006)

found no changes to analysts' forecast dispersion and forecast error, while Bailey *et al*. (2003) Agrawal *et al*. (2006) and Wang (2007) found that analysts forecast dispersion and forecast error increased following the regulations suggesting greater information asymmetry as a result of the regulation. Within the second domain, Kross and Suk (2012) found that the speed and magnitude of forecasts revisions following disclosure of public information by corporate management have increased significantly, suggesting that analysts rely more on public disclosure, given the unavailability of the private disclosure channel. However, recent survey evidence in Brown *et al*. (2015) reveals that the incentive to curry favour with managers is still influential to the work of analysts. Survey and interview respondents confirm that they still benefit from private conversations as it provides an opportunity to double-check their models, infer optimism about the firms' prospects through analysis of managers' body language which provides "colour" and "granularity" (p. 23). Moreover, the need to curry favour with corporate managers goes beyond access to information. As discussed in the next session, the incentive to obtain and maintain investment banking business is also a motivation. Thus, analysts' incentive to curry favour with firm management may be limited but not completely eliminated, despite recent regulatory pronouncements.

7.3.2 Investment banking and brokerage business

Anecdotal evidence documents several cases in which IB-analysts issued biased recommendations and forecasts in order to curry favour with corporate managers and obtain investment banking business. Table 25 lists accounts from the case against Merrill Lynch by the New York Attorney General, Eliot Spitzer in the wake of the dotcom crisis, in which analysts publicly issued positive recommendations which were contrary to their true beliefs about the company's prospects as revealed in internal email communications.

Table 25: Disparity between private and public ratings by analysts at Merrill Lynch

Company	Date	Analysts' email comments on the company	Published ratings
Excite@home	27/12/1999	We are neutral on the stock	2-1
Excite@home	03/06/2000	…Such a piece of crap	2-1
Infospace	20/10/2000	Piece of Junk	1-1
Internet Capital Group	06/10/2000	No hopeful news to relate…We see nothing that will turn around near term. The company needs to restructure its operations and raise additional cash and until it does that, there is nothing positive to say	2-1

Source: Adapted from Swedberg (2005)

Note: The tables was obtained from the affidavit of the New York Attorney General, Eliot Spitzer on 8 April 2000 against Merrill Lynch and it reveals the disparity between internal email comments about the companies and their published ratings. The ratings are based on Merrill Lynch's 5 point system with 1-Buy, 2-Accumulate, 3-Neutral, 4-Reduce and 5-Sell. For example, a 2-1 would mean accumulate in the short-run and buy in the long-run.

Additionally, it was observed that 6 out of the 8 independent analysts-firms following Enron's shares prior to its collapse issued "sell" recommendations before any of the other IB-analysts (Coffee, 2002 and Labhart, 2004) supporting the differential tendency to bias by IB-analysts and IND-analysts due to investment banking incentives.

Empirically, several studies have sought to examine the influence of investment banking business on analysts' research output. Usually, this involves observing differences in recommendations and forecasts issued by analysts who differ based on their perceived incentives to curry favour with management for investment banking business. Some studies focused on existing investment banking relationships between the companies and investment banking firms (distinguishing between affiliated vs unaffiliated analysts) while others examined both existing and potential relationships by splitting sell-side analysts based on the type of employers (e.g. investment banks, brokerage firms and independent research firms).

For the first group of studies, Lin and McNichols (1998), Michaely and Womack (1999), O'Brien *et al*. (2005) and Kolanski and Kothari (2008) provide some evidence which supports the claim that investment banking business motivates optimism among affiliated analysts. Based on the premise that certain corporate

events create more incentive for bias, some of these studies have focused on the effect of bias on analysts' earnings forecasts and recommendations in the context of seasoned equity offerings, initial public offerings and mergers and acquisitions. Lin and McNichols (1998) found that long term growth forecasts and stock recommendations are more optimistic for affiliated than unaffiliated analysts following seasoned equity offerings. Relatedly, Michaely and Womack (1999) examined the issue of bias by affiliated analysts within the context of initial public offerings. Their findings are supportive of Lin and McNichols (1998) as they found that affiliated analysts were more likely to issue positive recommendations following initial public offerings than other analysts. Moreover, stock returns following positive recommendations by affiliated analysts are lower than those of unaffiliated analysts, suggesting that investors acknowledge these biases and discount recommendations accordingly.

Kolanski and Kothari (2008) examined the behaviour of analysts affiliated with investment banks which advise acquirers and targets in mergers and acquisitions. They conjecture that analysts affiliated with investment banks which advise the acquirers are incentivised to be optimistic about the acquirers. The positive share price reaction is desirable for the acquiring firm management because it increases the chances of the deal being approved by shareholders and increases the "currency" of the acquisition in a stock deal. The IB-adviser also benefits because the share price increases the likelihood of acquisition and may potentially increase the fees received. On the other hand, analysts affiliated with IB-advisers of the target firm are incentivised to issue optimistic reports about the acquirer only after the exchange ratio in a stock deal is secured. They found, consistent with their hypotheses, that analysts affiliated with the acquirers are more likely to upgrade their recommendations within 90 days from the mergers and acquisition deal, while analysts affiliated with the target are likely to issue a positive recommendation for the acquirer only after the exchange ratio of the stock deal has been set. Their results support the claim that analysts, whose employers are involved in a merger and acquisition deal as advisors, strategically report about the target and acquirer to benefit corporate managers and increase the likelihood of securing the deal for their employers.

While the previous studies were event-driven and examined measures of bias using either earnings forecasts or recommendations, O'Brien *et al*. (2005) examined the influence of investment banking relationship on bias by observing the speed with which analysts convey unfavourable news. Their results support the conjecture that investment banking ties influence the timeliness with which analysts' respond to news. They show that affiliated analysts' downgrades from buy and hold recommendations are slower than those from unaffiliated analysts, whereas upgrades from hold recommendations are timelier from affiliated compared to unaffiliated analysts. The study offers interesting insight into how analysts' conflicts of interest affect various aspects of the research process, including the speed with which they downgrade/upgrade their recommendations.

The second stream of studies examines the influence of employers' business as a motivation for analysts' bias by investigating both existing and potential conflicts. Thus, rather than observe the difference between affiliated (who have existing relationship with firms) and unaffiliated analysts, they distinguish between analysts based on the type of firms which employ them or on the magnitude of the revenue received from each business. Early studies such as Dugar and Nathan (1995) observed that IB-analysts are relatively more optimistic in their earnings forecasts and recommendations than non IB-analysts. Their findings also suggest that investors are aware of this bias and respond adequately by discounting the recommendations issued by IB-analysts. Barber *et al*. (2007) found that buy stock recommendations of independent firms (independent research firms and brokerage firms) outperform those of their IB-counterparts, while hold and sell recommendations underperform those of IB-analysts. They attribute this to the loss of credibility in buy recommendations as a result of investment banking business.

Other studies failed to support greater optimism by IB-analysts relative to non-IB analysts. For instance, Jacob *et al*. (2008) found that earnings forecasts by IB-analysts are more accurate and less optimistic than those of non IB-analysts. Moreover, among IB-analysts, affiliated analysts were also found to be more accurate than unaffiliated analysts. They attribute this finding to several factors which might positively influence IB-analysts research quality: First, they argued that investment banks have greater resources to fund research, hence are able to

employ higher quality analysts, who are able to generate more accurate forecasts. Second, they suggest that the provision of investment banking services beyond research provides an information advantage for IB-analysts, given that they have greater access to management and are more informed about the prospects of the companies they follow. Relatedly, Gu and Xue (2008) split their sample between independent analysts (independent research firms and investment managers) and non-independent analysts (investment bankers and brokerage firms). They extend prior studies by distinguishing between *ex-ante* and *ex-post* measures of research quality. The *ex-ante* measure of research quality is based on the association between analysts' forecast errors and abnormal returns. They argued that forecasts which proxy for market expectations of earnings would result in forecast errors which have greater association with abnormal returns. On the other hand, *ex-post* measures of performance are forecast bias and accuracy. They provide evidence similar to Jacob *et al*. (2008) that IB-analysts are less optimistic and more accurate than their independent counterparts. However, using an *ex-ante* indicator of research quality, they found that IND-analysts produce superior forecasts compared to IB-analysts. Finally, they document a disciplining role of independent analysts as both *ex-ante* and *ex-post* measures of research quality are higher for IB-analysts when the stocks are also followed by an IND-analyst.

Rather than the IB/non-IB dichotomy, Cowen *et al*. (2006) examined differences in forecast bias across four different types of firms based on their sources of research funding. They classify analyst-firms into four groups: (1) full-service banks that fund research through underwriting and distributing new issues and trading activities; (2) syndicate banks that fund research through fees from distributing new issues and trading activities but not underwriting; (3) brokerage firms that fund research only through trading activities and do not provide any investment banking services such as distributing or underwriting new issues; and (4) pure research firms that do not provide either investment banking or brokerage services. Their findings suggest that forecasts of EPS, long-term growth and price target and recommendations by analysts employed by brokerage houses are the most optimistic while analysts employed by full-service investment banks were found to be the least optimistic. Hence, trading concerns rather than underwriting concerns may be the main incentive for bias. Further confirmation of the relative importance of trading in shaping analysts' research output is also documented in

Agrawal and Chen (2012). However, unlike the previous studies, they distinguished analysts-firms based on the fraction of revenue from different businesses and found greater optimism in long term growth forecasts for analyst-firms with greater revenue from brokerage businesses. In general, the empirical evidence thus far is mixed and is in part influenced by differences in research design.

Following recent regulation[38] aimed at curbing the perceived influence of analysts-firms' business (particularly investment banking business) on analysts' objectivity, Clarke *et al*. (2011) found that the stock recommendations of analysts whose employers' have no investment banking business have less information content and are more optimistic than IB-analysts, questioning the supposed superiority of IND-analysts. Although Chen and Chen (2009) document increasing positive association between recommendations and the relation between intrinsic value and stock prices following these regulations, Barniv *et al*. (2009) document that stock recommendations are explicable by other factors beyond that association. Moreover, Kadan *et al*. (2009) suggest that affiliated analysts are still reluctant to issue pessimistic recommendations. Hence, employer-related bias may not be completely eliminated by recent regulations.

7.3.3 Institutional investors

Conflict of interest arises when institutional investors have large positions in stocks covered by analysts. To avoid portfolio losses on which their performance is based and attractiveness of the funds to potential investors, fund managers may pressure analysts not to issue unfavourable recommendations until they have disposed of the stock. Thus, certain institutional investors may influence the type and timing of information disclosed by analysts in their favour. Firth *et al*. (2013) provide recent evidence indicating that analysts are subject to pressures from institutional investors. They found that analysts issue higher stock recommendations relative to consensus if the stock is held by a mutual fund client of their employer. This association is stronger, the higher the proportion of the stock held by the mutual fund client and the higher the brokerage commissions received from the mutual fund. They also found that the tendency to be optimistic

[38] These have been discussed extensively in Chapter 2

was less pronounced for firms with more mutual funds ownership, suggesting that concerns over reputational loss help curb the optimistic tendencies. This moderating role is also documented in Ljungqvist *et al*. (2007). They found that institutional investors have a moderating role on analysts' bias given their ability to influence career outcomes through analysts' rankings. Empirically, their results indicate that recommendations issued by analysts are relatively optimistic compared to consensus the greater the investment banking and brokerage incentive. However, this optimism is curtailed the higher the institutional ownership of the shares and the greater the reputational capital of the investment bank. Also, they found smaller forecast error when the stock is held by institutional owners suggesting that concerns about reputation improve analysts' performance. Finally, they found that analysts' negative reaction to bad news in the form of downgrades is quicker the greater the level of institutional ownership. The empirical evidence thus suggests that the presence of dominant institutional investors can motivate analysts' to bias their recommendations upward or it can inhibit bias, thus resulting in less optimism. This source of bias is less documented and hence not the focus of this study.

Considering the significant role that information intermediaries such as analysts' play in the financial markets, the issue of bias remains one of interest for various stakeholders. The evidence from prior empirical studies as discussed in this section provides interesting results but the mixed evidence and the focus on earnings forecasts and recommendations limits our understanding of several aspects of analysts' bias. There are several factors which motivate an extension of prior research on analysts' bias to an examination of the narrative content of their research reports. First, empirical evidence indicates that the narrative content of analysts' research has information content over and beyond the earnings forecasts and recommendations (see literature reviewed in section 7.2). This implies that the narrative content of analysts' reports is important to investors and impacts on their investment decisions, and as such, is an important element of the output of analysts' research activity. Second, survey evidence suggests that institutional investors are able to generate earnings forecasts and stock recommendations using their in-house models and are more reliant on analysts' for the additional industry analysis and insight detailed in their narratives (Imam, 2012). A former analyst explains this as follows:

> In the world of professional money management, most clients (called the "buy side") don't care at all about an analyst's recommendation or target price. Instead, they care about the data an analyst may have gathered, or the critical thinking that went into making a particular recommendation or earnings forecast. In other words, professional investors already have an investment thesis on a stock they are following, and they really don't look to sell-side analysts to give them bottom-line recommendations. Instead, they gather all kinds of varying viewpoints from competing analysts and consider which ideas, opinions and data points make the most sense.
>
> (Umiastowski, 2014)

Additionally, institutional investors rank written reports much higher than earnings forecasts, when asked to rate the importance they attribute to analysts' work (Institutional Investor, 2010). This calls for a focus into the narrative content of the reports rather than recommendations and forecasts as the narratives may actually be the single most valuable output of analysts' research activity for investors. Finally, research in social psychology has highlighted the importance of studying the features of people's communication. A widely held view in this field of study is that "how" people communicate contains important information beyond "what" is communicated and "words people use are diagnostic of their mental, social and even physical state", (Pennebaker *et al.*, 2003, p. 548). The implication is that a study of the linguistic features of analysts' reports can reveal additional information beyond the thematic content of analysts' reports and analysts who are incentivised to be optimistically biased may also strategically select a writing style that achieves their purpose. As previously highlighted, research has documented evidence of strategic use of linguistic features in corporate reporting, but this is yet to be extended to analysts' reporting. Hence the research objective of this part of the thesis is to examine whether analysts' incentive to produce optimistic research explains variation in the linguistic features of their reports. Incentive to bias is proxied by analysts' employer. As in previous research reviewed above, a distinction is made between analysts employed by investment banks and brokerage houses and analysts employed by independent research firms. Chapter 2 discusses the features of both types of analysts. The first step in addressing this research objective is to identify a theoretical framework which explains why analysts' incentives may influence production of optimistic research.

7.4 Theoretical framework on analysts' bias

Despite extensive research on analysts' bias as reviewed in the previous section, a comprehensive theoretical framework is still elusive. Hayward and Boeker (1998) used theory of *power* to explore the cases of conflict of interest derived from investment banking business. They focused on intra-organisational conflict between the equity research department and the corporate finance department within investment banks. They argue that a party's relative power and control of desired resources within the organisation will determine their actions. They posit that, given that corporate finance departments are one of the main sources of an investment banking firms' revenue while the equity research departments are mostly cost-centres, the interest of the equity research department in producing objective research is overpowered by the interest of the corporate finance department to attract and retain clients. They argue that more powerful groups are able to influence outcomes such that presence of corporate finance departments in analysts-firms will result in more positive stock recommendations. The analysis was based principally on the power play between departments in investment banks and did not considered other sources of analysts' conflict of interest such as corporate managers.

Generally, most studies of analysts' bias have relied on the concept of conflict of interest (Michaely and Womack, 1999). This is premised on the existence of a principal-agent relationship from an economic perspective. However, the actual fiduciary relationship between analysts and investors or other parties, which influence their research output has not been clearly defined. Fisch and Sale (2003) challenge the assumptions of a fiduciary relationship between analysts and investors given that analysts' conflict of interest derives principally from their role as agents for different principals. First, they act on behalf of their employers, whose interests may diverge from those of investors and corporate managers. Second they act on behalf of corporate managers in disseminating information about the company to the capital market, whose interest diverges from investors and, finally, they act on behalf of investors, both existing and potential, by providing the needed information for decision-making. Formally, analysts have a contractual relationship with their employers. Further, analysts are also influenced by corporate managers, given their dependency on them for information, access and investment banking business (Fogarty and Rogers, 2005)

and act in their interests (Brown *et al.*, 2015) despite the lack of any contractual obligations. Thus, a social perspective is adopted to explain analysts' motivation to "curry favour" with corporate managers. Francis *et al.* (2004) explains the influence of corporate managers as follows:

> "Analysts' fears of reprisal for issuing negative reports appear justified. According to a survey by Tempest Consultants sponsored by Reuters Holdings PLC, when asked how they would respond if an analyst issued a sell recommendation on their company, about one-third of company managers responded that they would not include the analysts' employer in future investment banking business and would reduce communications with the analyst and the analysts' access to management." – Francis *et al.* (2004)

Westphal and Clement (2008) used theories of social exchange and social norms of reciprocity to examine the exchange of favour between corporate managers and analysts. While the focus was on managerial influence on analysts, they provide an alternative perspective on understanding why analysts may be predisposed towards optimism. They find that managers of companies with negative financial performance are more likely to render personal and professional favours to analysts, which is reciprocated by lesser stock downgrades. Moreover, they find that analysts who downgrade stocks elicit "negative reciprocity" from managers who retaliate by reducing personal access and other forms of favour. This evidence supports the claims by survey participants in Brown *et al.* (2015) that analysts are concerned about pleasing corporate managers to avoid being denied corporate access despite the recent regulatory pronouncements and better explains why analysts are motivated to provide optimistic research output. From the managerial perspective, favour-rendering is a means of influencing analysts to issue reports consistent with their self-serving interest. For the analysts, favour-rendering in the form of optimistic research is a means of "currying favour" or ingratiating corporate managers on whom they depend on for rewards. Consequently, the social psychology theory of Ingratiation is relied on to explain analysts' optimism and formulate testable hypotheses.

Ingratiation is an impression management technique aimed at increasing one's attractiveness in order to get some rewards or avoid punishment. Pandey and Bohra (1984, p.382) define ingratiation as "motivated behaviour directed towards the goal of eliciting increased *attraction* from a particular person leading to a

favourable *reward* or decision". Similarly, Kumar and Beyerlein, (1991; p.619) define ingratiation as "a pattern of interpersonal influence behaviour which enhances one's *attractiveness* to others. Jones (1964, p.11) states that "the term ingratiation refers to a class of strategic behaviours illicitly designed to influence a particular other person concerning the *attractiveness* of one's personal qualities." Ingratiating is said to be illicit because it is directed toward objectives not contained in the implicit contract which underlies social interaction. Tedeschi and Melburg (1984, p. 37) refer to ingratiation as "a set of assertive tactics which have the purpose of gaining the *approbation* of an audience that controls significant *rewards* for the actor." Ingratiation is based on the assumption that people behave in order to attract reward rather than punishment. Hence, they act in accordance with expectations and demands of the settings that result in rewards. As stated in Gordon (1996, p.67), the primary goal of ingratiation is that of increased liking. Thus ingratiation occurs for self-benefit i.e., the ingratiator hopes to improve his own outcomes by biasing the target person in their favour. The four main components of this model as derived from these definitions is that there is an ingratiator, an ingratiatory behaviour, a target to which the behaviour is directed and an expected reward which the ingratiator hopes to obtain.

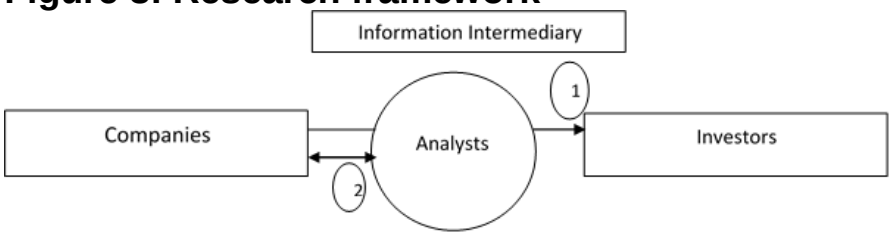
Ingratiation has been previously used to study organisational behaviour and evidence from prior research reveals that ingratiation techniques result in favourable outcomes for the ingratiator such as salary increases, favourable performance evaluations or career advancement (Westphal and Stern, 2006). Within the accounting literature, Robertson (2010) provides evidence that client-auditor relationships are often mediated through ingratiation as clients influences auditor judgement through ingratiation. In general, impression management is mostly concerned with managing impressions of others (Goffman, 1959) which may include enhancing personal attractiveness. Thus, ingratiation is a type of impression management technique which provides a useful framework for predicting how the relationship between analysts and corporate managers influence their research output. As identified in Schlenker (1980, p.171), ingratiation differs from other impression management strategies for the following reasons:

- Its main goal is to be attractive to others
- It places emphasis on pleasing others
- It is based on the expectation of obtaining personal rewards
- There is a strong possibility of dissonance between private beliefs and public statements.

Analysts are dependent on various parties (principally corporate management) for certain rewards. Hence, they are incentivised to act in a manner that improves their attractiveness to them in order to obtain these rewards. Evidence from prior empirical research on analysts' bias reviewed in section 7.3 of this Chapter suggests that analysts act to 'curry favour' with firm managers in order to obtain access to information or to obtain/maintain investment banking business (Ke and Yu, 2006; Brown *et al.*, 2015). The rewards that accrue to analysts from pleasing managers and generating business for their employers have also been previously documented. These include, better compensation, promotion, invitation to conference calls, access to managers through private phone calls, higher institutional ranking and employment at larger IB-firms (see, for instance, Hong and Kubick, 2003; Ke and Yu, 2006; Groysberg *et al.*, 2008; and Brown *et al.*, 2015).

Figure 7.1 provides a framework for illustrating the role of analysts and managerial influence on their objectivity.

Figure 8: Research framework



As illustrated in Figure 8, the traditional role of analysts is that of information intermediaries who mediate the conflict of interest between companies and investors by providing an 'objective' assessment of the firms' performance and prospects (Jensen and Meckling, 1976) - arrow 1 in Figure 8. However, managers who are charged with running the business, have self-serving interests and use a variety of impression management techniques to present their company performance in a positive light (Brennan *et al.* 2009). Their activities are

scrutinised by external monitors such as information intermediaries like analysts. Westphal and Clement (2008) argue that concerns about negative recommendations and downgrades motivate managers to influence analysts through provision or denial of personal and professional favours. Further, norms of reciprocity in social exchanges imply that analysts' respond through provision of optimistic research which has been extensively documented in the literature (see section 7.3). Hence, while analysts are not formally employed by the companies they report for, there exist a social contract between analysts and corporate managers which consists in the exchange of favours and influences analysts' intermediary role - arrow 2 in Figure 8.

7.4.1 Causes of ingratiation

Why would analysts seek to curry favour with managers? What motivates ingratiatory behaviours? Jones (1964) presents a theoretical framework for the study of ingratiatory behaviour and suggests that ingratiation behaviours are motivated, among others, in circumstances in which a target person controls a scarce or valuable resource which the ingratiator hopes to acquire at a minimum cost to himself. This expected benefit is presumed to exceed the level that would normally be obtained in the course of legitimate social exchange. Furthermore, Kacmar *et al*. (2004, p.310) suggest that "two major antecedents of ingratiatory behaviours are the *situation* in which the ingratiator finds himself and his or her *disposition*".

In this study, the employment structure of analysts is posited as being a situational factor that motivates ingratiation. While analysts in general depend on managers for access to corporate information, the incentive to curry favour with managers through optimistic research is higher for IB-analysts. For the IB-analysts, the rewards from currying favour with managers go beyond the need to access information and are linked to other businesses offered by their employers. First, it improves the likelihood of securing investment banking business. Second, it enable analysts to provide corporate access to the buy-side (Soltes, 2014) which in turn translates into higher broker votes[39], institutional ranking and consequently

[39] Broker votes are ratings of the value of broker's research services and are produced by institutional investors. They influence the level of trade allocated to a brokerage firm and

better career prospects (Brown *et al.*, 2015). However, for the IND-analysts, the incentive to curry favour with corporate managers for investment banking business and corporate access which influences brokerage commissions, is non-existent given the lack of investment banking or brokerage business in such firms.

7.4.2 Ingratiatory tactics

Ingratiation can take any form by which interpersonal attraction may be solicited (Jones, 1964). These techniques will vary across contexts and the preferences of the audiences. Prior studies have usually categorised ingratiation tactics into four classes as follows: opinion conformity, other enhancements, self –promotion and favour-rendering.

Other-enhancement:
This represents a set of tactics aimed at enhancing the targets' strengths and virtues, through flattery and praise. For instance, Fogarty and Rogers (2005) found that words related to praise and accomplishment were used by analysts to discuss managers and corporate plans in general.

Opinion conformity:
These tactics include agreement with the targets' opinions, attempts to articulate the position presumed to be held by the target or behaviour imitation and identification. The proposition of this class of tactics is that people like others whose values and beliefs appear to be similar to their own. For instance, Feng and Mcvay (2010) examined the extent to which analysts conform to the earnings guidance provided by managers. They found that only analysts with incentive to curry favour with managers overweigh managerial guidance when forecasting earnings.

Self-Presentation:
This involves the explicit presentation or description of one's own attributes to increase the likelihood of being judged attractive. In other words, ingratiation takes the form of actively presenting oneself along the lines of the target person's suggested ideal.

---

therefore their commission from trading and influences analysts' compensation. (Maber *et al*. 2014)

Favour-rendering:
These tactics involve the offering to perform (or actually performing) a task for the benefit of the target. In other words, the ingratiator actually does a favour for the target in the hope of being liked and the expectation that the favour will be returned in some form of reward for the ingratiator. The favour rendered by the ingratiator creates a debt owed by the target due to the norms of reciprocity (Gertsen 2009; Kacmar *et al.*, 2004). Westphal and Clement (2008) provide evidence of this in the context of analysts and corporate managers. Similar to their study, the ingratiation tactic examined in this thesis is that of favour-rendering. It is assumed that the analysts' optimistic research is a favour rendered to corporate managers, who in-turn reward this favour with personal and professional favours such as increased opportunities for private access to managers.

In prior literature on analysts' optimistic bias reviewed in section 7.3, analysts' stock recommendations and earnings forecasts have been the instrument of optimism examined. In the present study, the examination of optimism is extended to the narrative content of analysts' reports. There are several ways in which the linguistic features of the narratives in analysts' reports could be manipulated to present an optimistic view of companies. However, this research focuses on tone and readability as it seeks to extend prior studies (discussed in section 7.2) which document the information content of these features in analysts' reports. The following extracts illustrate how analysts may bias their reports to present a positive view of the stocks they follow:

> "Adjusted Q2/11 EPS of $0.32 was roughly in line with our estimate of $0.33 and matched consensus of $0.32."

(Source: RBC research on Alcoa, July 11, 2011)

> "Excluding one-time items, Alcoa earned $0.32 per share in 2Q11, matching the consensus forecast and exceeding our estimate of $0.31."

(Source: Argus research on Alcoa, July 12, 2011)

The extracts are from reports by an IND-analysts and IB-analysts on the same company. In the first case, a $0.01 EPS miss was described as being "roughly in line", while in the second case, a $0.01 EPS beat was described as "exceeding".

The former fails to identify an EPS miss which might be viewed negatively while the latter stresses the EPS beat with a strong positive word. As revealed in the illustrative extracts the same information could be presented in different ways that convey additional information beyond the hard quantitative information. The psychology literature suggests that word usage is reflective of the writers' psychological state and motivation (Pennebaker *et al.*, 2003; Newman *et al.*, 2003). The assumption in the present study is that linguistic features are reflective of analysts' psychological processes and, therefore, provide information about analysts' bias beyond the recommendations.

7.5 Chapter summary

The purpose of this chapter was to review the literature on the linguistic features of analysts' reports and the literature on analysts' bias while identifying the gaps in prior research and introducing the research objective and the theoretical framework.

The chapter commenced with a review of the literature on the linguistic features of analysts' reports. The review revealed that only a few studies have examined the linguistic features of analysts' reports with most studies focusing on the information content of the tone and readability of the reports. This was contrasted with the extensive literature on corporate reports that indicate the strategic use of a variety of linguistic techniques by managers seeking to present a positive view of their company's performance.

Thus, the research objective of this study emerged as it aims to extend the findings from corporate impression management studies to the context of sell-side analysts by examining whether the variation in the linguistic features of tone and readability of analysts' reports is explained by analysts' incentive to produce optimistic research, which is proxied by analysts' employment structure.

The extensive literature on analysts' bias was further reviewed which explained the various factors that incentivises analysts' to issue optimistic recommendation and forecasts and provide empirical evidence both pre and post-regulatory pronouncements aimed at enhancing analysts' objectivity. Empirical evidence from this literature was mixed and limited to study of recommendations and

earnings' forecasts, which further justifies the need to examine other analysts' output such as their reports.

Finally, the impression management theory of ingratiation was introduced as providing a useful framework for understanding analysts' predisposition to bias and is relied on to formulate testable hypotheses in the next chapter.

Chapter 8: Hypotheses development II

8.1 Introduction

In the previous chapter, the main argument put forward and on which hypotheses in this chapter are based, is that IB-analysts are more incentivised to curry favour (or ingratiate) corporate managers than IND-analysts. Hence, IB-analysts are more predisposed to produce biased research output. Rationale for this conjecture has been presented and discussed, including a review of extant empirical evidence. As previously highlighted, existing studies which examine analysts' predisposition to optimism have been based on examination of summary measures of analysts' research output such as recommendation and earnings' forecasts. This present study is based on the linguistic features of analysts' narratives. Of the variety of linguistic features previously studied in the context of strategic corporate reporting, this study focuses on the examination of tone and readability of analysts' reports. Hypotheses 1 and 2 test the difference in tone and readability between IB and IND-analysts' reports in the sample. This is presented and discussed in section 8.2 and 8.3. Hypotheses 3 and 4 are presented in section 8.4 and are aimed at further test of strategic reporting by examining association between the linguistic features and recommendation levels. Section 8.5 highlight the models used to test the set hypotheses and section 8.6 concludes the chapter.

8.2 Differences in level of optimism in IB and IND-analysts' reports.

The first set of hypotheses examine the paired differences in the tone of IB and IND-analysts' reports. Westphal and Clement (2008) show that optimistic research output is a part of a "favour-rendering" exchange between analysts and corporate managers as analysts curry favour through optimistic research. Further, Ingratiatory theory suggests that ingratiatory behaviour is aroused among others by the importance of the reward to the ingratiator (Jones, 1964). As previously discussed in Chapter 7, the incentives to produce optimistic research and the rewards that derive from this are greater for IB-analysts compared to IND-analysts. For instance, access to managers enable IB-analysts to also gain the opportunity to introduce their buy-side client to corporate managers (Soltes, 2014). This is valued by the buy-side and translates into higher ranking for the analysts (Brown

et al., 2015). Thus, IB-analysts are alleged to be more incentivised to issue optimistic research as a means of currying favour (i.e. ingratiating) with corporate managers. Several studies have examined the differences between IB and IND-analysts' research output to test the conjecture that IB-analysts are more predisposed to optimism than IND-analysts (see section 7.3). However, these studies have largely focused on earnings forecasts and recommendations (e.g. Gu and Xue 2008; Jacob *et al.*, 2008; Clarke *et al.*, 2011). Moreover, results till date have been equivocal and despite recent regulatory pronouncements, analysts still appear to curry favour with corporate managers (Brown *et al*. 2015). Furthermore, the perceived objectivity of IND-analysts (given the absence of investment banking and brokerage business), have also influenced the actions of regulators. For instance, one of the provisions of the 2003 Global Settlement Act was the mandatory requirement that IB-firms supplement research by their analysts with reports from IND-analysts. The specific terms with regard to independent research were summarised by the SEC as follows:

> For a five-year period, each of the firms will be required to contract with no fewer than **three independent research firms** and will make available the independent research to the firm's customers. Firms will notify customers of the availability of independent research on customer account statements, on the first page of research reports, and on the firm's website. An independent consultant for each firm will have final authority to procure independent research, and will report annually to regulators concerning the research procured. Payments for independent research will total $432.5 million
>
> *(SEC fact sheet on Global Analyst Research Settlements)*

The implicit assumption of this provision is that IND-reports are less biased than those issued by IB-analysts. Consequently, it is expected that IB-reports will be more optimistic than IND-reports. Unlike previous studies, optimism is measured by the tone of the reports rather than the recommendations or earnings' forecasts.

***H_{1A}**: The tone of IB-analysts' reports is more optimistic than those of IND-analysts' reports.*

H_{1A} aims to test the extent to which IB-analysts are relatively more optimistic compared to IND-analysts who are perceived as being less biased. A second test

of analysts' optimistic bias is the extent to which optimism is founded on the company's performance.

The level of financial performance is perceived as an important factor that encourages managers to engage in impression management. Clatworthy and Jones (2006) found that the Chairman's statements of poor performing firms contains linguistic features which are indicative of impression management such as less self-referencing pronouns and more passive sentences. Courtis (2004) and Li (2008) found that corporate reports are harder to read when earnings are lower or less persistent. Similarly, Cho *et al*. (2010) show that firms with poor environmental performance issue environmental disclosures which are more optimistic and less certain. The premise as argued by Merkl-Davies and Brennan (2011, p.425) is that "management engages in impression management in anticipation of an evaluation of its actions and decisions primarily by shareholders which serves to counteract undesirable consequences." Thus, in the face of negative organisational performance, managers have greater incentive to engage in impression management to counteract undesirable consequences from investors to whom they are accountable (Westphal and Clement, 2008).

While the accountability argument is not descriptive of analysts' motivation as previously argued, the context of negative organisational outcome increase the value of analysts' optimism for the benefit of corporate managers. Jones (1964) argues that the ability of the desired goal to arouse ingratiatory behaviour is dependent on "*the ingratiators subjectively assessed probability that the ingratiatory behaviour will yield the desired reward."* Westphal and Clement (2008) provide relevant empirical evidence as they show that the announcement of negative earnings surprises increase the tendency of corporate executives to engage in social influence behaviour in the form of personal and professional favour rendering to analysts and analysts are less likely to downgrade their recommendations for such firms, as a result of such favours.

It is thus expected that if optimism is motivated by the incentive to ingratiate corporate managers, the tone of the reports will be less associated with firm fundamentals (such as financial performance). Penman (2003) commented that this practice was widespread during the dotcom bubble as optimistic analysts generally ignored the financial losses reported by most dot.com companies in

favour of other metrics such as 'number of clicks' per page on a website. In addition, Beunza and Garud (2007) corroborate this observation with an empirical study of the content of analysts' reports. Hence, the following hypothesis tests the extent to which the tone of reports issued by analysts who are more incentivised to issue optimistic research output (IB-analysts) is associated with financial performance.

H_{1B}: The tone of IB-analysts' reports is less associated with financial performance than the tone of IND-analysts' reports.

8.3 Differences in level of readability in IB and IND-analysts' reports.

The second linguistic feature examined in this study is the readability of analysts' reports. The level of readability has been previously associated with reporting bias within extant literature on corporate reporting (Courtis, 2004). It is argued that manipulation of reading ease is a means used by managers to obfuscate negative organisational outcome (Merkl-Davies and Brennan, 2007; Brennan and Merkl-Davies, 2013). Unlike for corporate managers, whose accountability relationship with investors influences their reporting choices (Merkl-Davies and Brennan 2011), analysts' incentives to produce optimistic research is hypothesised to influence the linguistic features of their reports such as readability. Optimism may be achieved through a variety of means which may include concealment of negative performance through more complex reporting. Thus, for analysts predisposed to optimism, less readable reports may be issued to conceal negative news or mask optimism. Chen and Chen (2009) admit that the concern in relation to analysts' bias is that recommendations and forecasts issued by analysts are not just optimistic but rather they are not reflective of analysts' true beliefs regarding the companies. This view accords with the concept of deception which "entails messages and information knowingly transmitted to create a false conclusion" (Zhou *et al.*, 2004, p.81) or involves control of the information contained in messages which convey a meaning that departs from the truth that communicators know (Buller and Burgoon, 1996, p.205). This is also consistent with the ingratiation literature which posits that, in ingratiatory settings, there is a "strong possibility that a discrepancy exists between privately held beliefs and publicly endorsed statements" (Schlenker, 1980). The case of Merrill Lynch analysts presented in Table 25 provides a practical example of how analysts may provide a public recommendation for companies which is at variance with their privately held beliefs. Thus, bad news may be concealed in the narratives which accompany any positive recommendations through manipulation of the readability of the reports, among other strategic reporting techniques[40]. Bushee (*et al.* 2014, p.1)

[40] For instance, Beunza and Garud (2007) provide example of selectivity bias in analysts' report. They examined the reports issued on Amazon during the dot.com bubble and showed that analysts' which issued more positive recommendations, focused on assessment of revenues while analysts issuing more negative recommendation of the same firm focused on assessment of earnings. Thus, analysts' predisposed to issuing optimistic research may

suggest that "whether complex language obfuscates information or conveys information depends on the disclosure-relative incentives of the source of complexity". Given IB-analysts greater incentive to optimistic bias, it is expected that the readability of their reports will be lower compared to IND-analysts.

What is more, similar to managerial incentive to manipulate reading ease to obfuscate negative financial performance (Li, 2008), incentive to curry favour with corporate managers is expected to result in more complex reports in the face of bad news. This pattern should be observable for IB-analysts given previously advanced argument that IB-analysts have a greater incentive to curry favour with corporate managers. Hence, it is expected that readability of their reports will be lower when financial performance is negative. Hypotheses H_{2A} test whether the readability of IB-analysts' report is greater than IND-reports for poor performing firms:

***H_{2A}**: IB-analysts' reports are less readable compared to IND-analysts' reports for the same company*

H_{2A} compares the readability of IB and IND reports conditioned on the news being negative. In addition, obfuscation is further tested by examination of the association between readability levels and financial performance and whether this association is different between IB and IND analysts. It is expected that if analysts varying readability level is indicative of bias, then readability should be significantly associated with financial performance as documented in previous studies on corporate reports. Hence,

***H_{2A}*:** The readability of IB reports is more associated with financial performance than the readability of IND report.

carefully select information items that signal positive news while ignoring negative information.

8.4 Strategic reporting

As a further test of whether variation in the content of analysts' reports is reflective of strategic reporting, the association between relative optimism in recommendation and differences in tone and readability between IB and IND-analysts' reports is examined and the association between the level of optimism as measured by tone and the level of readability is also examined. This is consistent with the argument advanced in Malmendier and Shantikumar (2014) that when optimism is motivated by incentives, bias is not limited to recommendations. While that study examines joint bias in recommendation and earnings' forecasts, this thesis extends this to the context of narratives which accompany the recommendations and examines whether relative optimism in recommendation is associated with relative optimism in the tone or accompanied by more complex reporting.

For the first set of tests, relative optimistic recommendation is defined as a situation in which for a given company, a positive recommendation by an IB-analyst is accompanied by a neutral or negative recommendation by an IND-analyst or vice-versa. Given the premise that incentive to curry favour with corporate managers is higher for IB than IND-analysts, it is anticipated that when IB-analysts issue more positive recommendations compared to IND-analysts following the same companies, the tone and complexity of the reports will also be higher. On the other hand, when recommendations by IND-analysts are more positive, the difference in tone and complexity will not be significantly different to those of IB-analysts. In other words, hypotheses stated below test the extent to which relatively more optimistic recommendations are associated with strategic reporting. It is expected that strategic reporting will accompany IB-recommendations as opposed to IND-recommendations.

***H_{3A}**: When recommendations by IB-analysts are more positive than IND-analysts, the tone of IB-analysts' reports is more optimistic than those of IND analysts' reports for the same companies.*

***H_{3B}**: When recommendations by IND-analysts are more positive than IB analysts, the tone of IND-analysts' reports is not significantly different from IB-analysts' reports for the same companies.*

***H_{3C}**: When recommendations by IB-analysts are more positive than IND-analysts, IB-analysts' reports are less readable compared to IND-reports for the same companies.*

***H_{3D}**: When recommendations by IND-analysts are more positive than IB-analysts, the readability of IND-reports is not significantly different from IB-analysts' reports for the same companies.*

The second sets of tests of analysts' strategic reporting is the examination of the association between level of optimism and readability. Following the argument in Malmendier and Shantikumar (2014), it is anticipated that analysts predisposed to optimism will accompany pessimistic reports with less readable writing to obfuscate the pessimism and accompany optimistic reports with more readable writing. Hence, the association between tone and readability is tested as evidence of strategic reporting and given IB-analysts' incentives, it is anticipated that this will be a stronger pattern for IB-reports compared to IND-reports. Hence:

H_{4A}: The readability of IB-analysts' reports is positively associated with the tone of their reports.

H_{4B}: The readability of IND-analysts' reports is not associated with the tone of their reports.

8.5 Testing the hypotheses

8.5.1 Statistical tests

The analysis is aimed at examining whether the linguistic features of analysts' reports (tone and readability) varies based on analysts' incentives to produce optimistic research. Distinction is made between IB-analysts and IND-analysts. Two types of analysis are conducted to test the hypotheses. Both paired sample tests and multivariate regression analysis are used to test the stated hypotheses.

Matched sample analysis is based on paired sample t-tests of difference between means and Wilcoxon signed rank test of difference in median. Further, a conditional logit regression model was estimated to examine the association between IB vs IND reports and the tone and readability of the reports. The basic model takes the form

$$Logit\ (p_i) = Log\ \left(\frac{p_i}{1- p_i}\right) = \ \alpha_0 + \ \beta_1 Tone_{j,i} + \beta_2 Bog_{j,i} + \beta_3 POS_{j,i} + \ \varepsilon_{ij} \qquad (2)$$

Where:

α_0 : Constant term

p_i: Probability that report (i) is issued by an IB-analyst, given the values of the explanatory variables.

β_{1-3}: Vector of coefficients on tone, readability and recommendations.

Tone : $\frac{(Positive\ words - Negative\ words)}{Total\ word\ count}$; As discussed in Chapter 9, this is based on a combined word list of Henry (2008) and Loughran and McDonald (2011).

Bog: *Bog index* which measures the readability of analysts' report in this study. Details are discussed in Chapter 9.

POS: Dummy variable which takes the value of 1 for positive recommendations and 0 for negative and neutral recommendations.

ε_{ij} Error term

The above model is estimated based on grouped data to control for the effect of matching as advocated in Cram *et al*. (2009). This is also referred to as conditional logistic model as the likelihood of p_i is relative to each group, where groups represent the case-control. In this study, the grouping variable is the ticker symbol

representing the company. Each group consists of two reports, one by an IB-analyst and another by and IND-analyst on the same company.

Multivariate analysis consists of test of association between the linguistic features and financial performance, while controlling for other firm characteristics. The estimated model is:

$$LF_{ij} = \alpha_0 + \boldsymbol{\alpha X}_i + \beta\, POS + \varepsilon_{ij} \quad (3)$$

LF_{ij}:	Tone/readability of report *i* for firm *j*.
α_0 :	Constant term
α	Vector of coefficients on company characteristics
X_i	Vector of company characteristics
Company characteristics :	Variables that measure company characteristics: Percentage change in EPS (P_EPS), Earnings volatility (Risk), Firm size (Size), Sales growth (Growth), Market-to-book ratio (MTB), Intangibles (Intan) financial leverage (Leverage), Industry (LTC).
β	Vector of coefficients on recommendations
POS:	Dummy variable which takes the value of 1 for positive recommendations and 0 for negative and neutral recommendations
ε_{ij}	Error term

8.5.2 Variables

The dependent variables are the linguistic features of tone and readability. Chapter 9 provides detailed explanation of the choice of these measures. The independent variables have been previously described in Chapter 4 of this thesis.

8.6 Chapter summary

This chapter presents the development of hypotheses used to address the second research objective of this thesis, i.e. the examination of the factors which explain variation of the tone and readability of analysts' reports. First, it presents and discusses hypotheses aimed at testing the differences in the tone and readability of reports issued by IB and IND-analysts. Second, hypotheses which test the association between the linguistic features and recommendation levels were formulated. Third, hypotheses aimed at examining the association between the linguistic features were also developed.

Finally, the statistical model used to test the hypotheses was presented.

Chapter 9: Objective 2 research design

9.1 Introduction

This chapter sets out the methods used to measure linguistic features for the second research objective of this thesis. Broader issues relating to the research design (such as research methodology, sample selection and composition and choice of independent variables) have been previously discussed in Chapter 4. Therefore, this chapter focuses principally on the specific issues relating to the second research objective. It begins by reviewing the linguistic approaches previously used to investigate bias within corporate disclosure studies with the aim of identifying the most suitable for this research. This review is contained in Section 9.2 while Section 9.3 describes the approach used to quantity the measure of tone for this study. Section 9.4 discusses similar issues for the measure of readability. Section 9.5 concludes.

9.2 Measuring linguistic features in prior accounting research

As previously noted, this study examines analysts' reports using content analysis. For the purpose of the second research objective which is concerned with the linguistic features rather than the types of information contained in analysts' reports, a syntactic content analysis is used. Merkl-Davies (2007, p.132) refers to this type of content analysis as involving the investigation of the linguistic aspects of texts such as readability. As stated in Merkl-Davies (2007, p.136), "research in corpus linguistics and recent advances in computational linguistics provide the building block for a syntactic content analysis." This stream of research is relied on for the study of the linguistic features identified previously, i.e. tone and readability.

Analysis of the linguistic features of narratives could take various forms. In accounting research, this can take an objectivist approach involving measurement of linguistic features using quantitative scores through pre-defined rigid procedures or it may take the form of an interpretative analysis which involves the researcher's subjective assessment of various aspects of the text. The former is referred to as content analysis while the latter comprises interpretative data analysis approaches such as discourse analysis, narrative analysis etc. These distinctions are explained in detail in Merkl-Davies *et al.* (2014) and Beattie (2014) and have been previously discussed in Chapter 4. As stated in that chapter, the method of analysis adopted in this thesis will take the form of quantitative content analysis and is consistent with the positivist epistemological and ontological stance of the overall research.

Within the stream of literature that examines the linguistic features of accounting narratives, analysis may be computerised or conducted manually. Moreover, computerised approaches may be based on a pre-specified wordlist or probabilistic algorithms (also referred to as machine learning). Table 26

presents a summary of these methods, their features and examples of previous studies in which they have been used to study the content of analysts' reports.

Table 26: Approaches to analysis of linguistic features

	Manual	Computer based	
		Dictionary based	**Machine learning**
Description	Based on human coding	Use of software programmes to search and categorise words into different categories based on pre-defined wordlist.	Use of computer programme to inductively classify test based on statistical properties after prior training.
Examples	Breton and Taffler (2001), Asquith *et al*. (2005) and Abhayawansa and Guthrie (2012)	Mokoaleli-Mokotelli *et al*. (2009); Twedt and Rees (2012)	Huang *et al*. (2014).
Context	Considers the context of word usage	Treat narrative as a "bag of words", doesn't consider the context.	Context considered at training stage.
Reliability	Less reliable of the three approaches as subject to human error and bias.	Highly reliable as completely dependent on computer programmes.	Reliability is enhanced through the use of computer algorithm but pilot analysis is subjective and influences the main analysis.
Validity	Validity is enhanced as context of word use is considered, so limits type I and type II error.	Validity is based on the extent to which lists captures the underlying feature of interest. It is enhanced through use of context specific dictionaries.	Validity is enhanced through the manual analysis at the pilot stage as context of word usage is considered and influences main analysis.

9.2.1 Manual analysis

Several studies examine linguistic features of accounting narratives such as tone using a manual content analysis approach. This involves analysis and classification of portions of text into pre-specified categories following human devised coding rules. For instance, Clatworthy and Jones (2003, 2006) manually examined several linguistic features of the chairman's statements of a sample of 100 UK firms including tone, self-references and passivity using a manual content analysis approach. Similarly, Asquith *et al*. (2005) devised a manual coding scheme to

extract the tone (which they refer to as *sentiment*) from a sample of analysts' reports. The measure (denoted as 'strength of argument') is computed as the sum of positive remarks less the negative remarks about several themes. A further example of manual coding of tone was offered by Abhayawansa and Guthrie (2012).

An advantage of this method is that the context in which words are used is generally considered. This limits the chances of false hits (type I error). Moreover, by manually reading the text, researchers are also able to discern subtleties in the tone of the message that may not be captured by generic word search (Schleicher and Walker, 2010). Thus, the chances of missing relevant tonal statements (i.e. a type II error) are also limited. Consequently, if coding is based on a rigorous, well-developed coding scheme, validity is enhanced using this method compared to computerised approaches.

Additionally, the use of this method facilitates joint coding of linguistic features and the underlying topic (if this is relevant for the research question examined). For instance, Abhayawansa and Guthrie (2012) measured tone of intellectual capital (IC, hereafter) statements in a set of analyst' reports to test whether they strategically report on IC.

There are also several limitations associated with such manual analysis of linguistic features. First, its labour intensive nature implies that coding can be time-consuming, resulting in small sample sizes which may not support statistical analysis or enhance generalisability of findings. Secondly, manual coding can be rather subjective. Although, these limitations can be mitigated by the use of well-written coding schemes, there is always a measure of subjectivity which threatens the replicability of the study and limits follow-up studies (Li, 2010).

9.2.2 Computerised analysis based on the use of wordlist

This method involves searching a piece of narrative for the occurrence of a list of keywords, classified into specific categories. Often, this entails the use of a text analysis program which uses a "mapping" algorithm to assign words to different categories (Li, 2010). Linguistic features are measured based on the frequency of occurrences of words in related categories. The output of the analysis is a

numerical figure which scores the text based on word frequency counts. Thus, this approach is also referred to as the "bag-of-words model" (Kearney and Liu, 2014, p.5).

Examples of the use of this method to study analysts' reports include Kothari *et al*. (2009), Mokoaleli and Mokoteli *et al*. (2009) and Twedt and Rees (2012). Given the reliance on the use of computer programs, this method enables study of large samples of documents, which enhances the power of empirical tests and enhances the generalisability of findings. This is particularly relevant for this study which aims to test the association between linguistic features of analysts' reports and other company and analysts' characteristics. The objectivity of the analysis enhances replicability, which in turns encourages follow-up studies and allows comparisons of findings across different studies.

A key limitation of this approach is that the context in which keywords are used is not generally considered. This may threaten the validity of the measures obtained from the analysis. It is argued that the meaning of words can only be discerned in the context in which it is used (Milne and Adler, 1999). This is particularly relevant for cases of homonyms in which words spelt alike convey different meanings, e.g. "race" as in running and "race" as in descent. Moreover, in a situation in which the writing style is subtle such that the tone of a text is not easily identified by a single word, the chances of type II error might be increased.

Several computer programmes, with in-built wordlists, have been developed to analyse text in this manner. The most widely used software programmes for linguistic analysis (within the field of accounting) are the *General Inquirer (GI)*, developed by Harvard psychologist Philip Stone, the *Linguistic Inquiry and Word Count (LIWC)* by University of Texas psychologists James Pennebaker and *Diction* by Professor Hart of the University of Texas at Austin and Professor Craig Carroll at the Stern School of Business at New York University. However, given the developmental background in psychology and politics, the in-built dictionaries in these programmes have been found to be somewhat unsuitable for analysis of financial documents (as argued in Loughran and McDonald, 2011). Hence, another criticism advanced against this method is the validity of the measures derived from these programmes (Li, 2010). For instance, about 3/4th of the negative wordlist was found to have no negative connotation within financial and business

communication (Loughran and McDonald, 2011). To address this limitation, Loughran and McDonald (2011) introduced a wordlist that is more pertinent to the business environment and has been found to produce a superior classification of the financial sentiment in accounting narratives (Feldman *et al.,* 2010). This was a significant contribution to the literature as, prior to their study, researchers have relied on the generic lists such as the *GI* list to analyse financial documents including analysts' reports (for instance Kothari *et al.*, 2009). Subsequent studies have increasingly relied on context-specific dictionaries such as that developed by Loughran and McDonald (2011) and Henry (2008). This is possible because the textual analysis programmes generally enable the use of user-defined dictionaries. Thus, customised wordlists are increasingly being used to enhance validity. The challenge for researchers is to develop context-specific wordlists to enhance the validity of the data analysis.

9.2.3 Machine learning

This approach falls into the category of artificial intelligence systems as identified in Deffner (1986, cited in Morris 1994). It is based on the use of probabilities to make inferences about the characteristics of the text examined. Deriving from the field of computational linguistics[41], it applies statistical properties to determine probabilities of words pertaining to a specific category using computer algorithms. The process, as explained in Kearney and Liu (2014) is an inductive approach in which the main analysis is preceded by a manually coded pilot study which is used in training the system prior to the main analysis. Based on the classifications of words at the training stage, computer algorithms are formulated and used to analyse the main sample.

Li (2010) provides an example of the application of this method to accounting narratives. He used the naïve Bayesian learning approach to identify the tone of forward-looking statements in a sample of MD&A sections of 10-K and 10-Q filings. Huang *et al.* (2014) also used this method to classify a sample of analysts' reports into positive, neutral and negative tone categories.

[41] This is an interdisciplinary field which applies statistical rules to natural language processing.

With this method, context-specific analysis is enabled (Huang *et al.*, 2014). In other words, given that classifications of text portions into various categories is not based on a standard list of words but rather the outcome of the manual analysis at the training stage, the difference in word usage across different areas is captured which enhances validity. Moreover, the reliance on computational analysis using computer algorithms facilitates cost and time efficient analysis of large samples.

The use of computer programming to extract measures of different linguistic features gives an impression of objectivity. However, unlike the individual word count system, replicability might not be achieved if the coding at the pre-testing stage is not reliable. For instance, Li (2010) pre-coded 30,000 forward looking statements from the MD&A sections of 10-K and 10-Q filings to train the computer program before actual coding. He used 15 research assistants to achieve this but fails to report on whether there was an assessment of inter-rater agreement or whether there was a coding rule developed to classify the sentences. In the absence of such schemes, subjective assessment and human errors in coding at the pilot stage are fed into the main analysis which may threaten the reliability of the overall coding. Moreover, as highlighted in Huang *et al*. (2014), this approach, like the dictionary-based approach, ignores the internal structure of the words in the text. Therefore the context of word usage is only partially accounted for. Finally, this method is relatively less established in the literature and difficult to understand without prior background in computational linguistics.

9.3 Method used in this study

To address the second research question of this study, the computerised word-count system is used. This was deemed most appropriate for several reasons. First, compared to manual coding, it enables coding of a larger sample of text within a shorter time frame, resulting in empirical tests with higher power. This is a desirable feature because the goal of the textual analysis in this part of the study is to obtain measures that can be associated with other firm and analysts' characteristics for the purpose of making inferences. Second, this approach is more established in the accounting literature and therefore allows comparisons with other studies in this area. Moreover, the method yields highly reliable results. A key limitation of this approach, as above-mentioned, is the validity of the measures derived from the computer analysis. To enhance validity, context-specific wordlists are used for the main analysis, which is also preceded by a pilot test of the relative effectiveness of several lists in capturing linguistic features of tone.

9.3.1 Steps in data analysis

The steps undertaken to measure the linguistic features of the analysts' reports are presented in Table 27. The first step involves data preparation, which is followed by the selection of an appropriate wordlist for the analysis and data collection.

Table 27: Steps in data analysis

Steps	Section of Chapter 9
Data preparation	9.3.1.1
Wordlist selection	9.3.1.2
Data collection	9.3.1.3

9.3.1.1 Data preparation

The first step in the analysis involved obtaining and preparing the analysts' reports. This principally involved downloading the reports from the *Investext* Plus database, copying the narrative content into a word file and uploading this on Nvivo. Notes regarding data preparation have already been discussed in Chapter 4 of the thesis and contain details of the criteria for selecting reports, sample composition and section of the reports analysed.

9.3.1.2 Wordlist and sensitivity analysis

Wordlists choice

Following data preparation, the next step was to identify both the wordlist and the textual analysis software to be used. Within extant literature, some authors have used the in-built dictionaries of software programmes as in Mokoaleli-Mokotelli *et al*. (2009) while others measure linguistic features using multiple wordlists (e.g. Demers and Vega, 2011; Rogers *et al*., 2011; and Twedt and Rees, 2012). Among authors that use multiple wordlists, some use a single wordlist for the main analysis and other wordlists for sensitivity checks (e.g. Twedt and Rees, 2012) while other authors preferred to compute a score which aggregates the measures obtained from multiple wordlists for the main analysis. For instance, Rogers *et al*. (2011) used principal component analysis to construct a single variable for tone based on the wordlists from *Diction*, Henry (2008) and Loughran and McDonald (2011). They also reported a 25% overlap between the Henry (2008) and *Diction* wordlists, while Loughran and McDonald's (2011) negative wordlists contain a 59% overlap with Henry's (2008) and a 30% overlap with *Diction*'s. This suggests that a combined wordlist may provide a useful alternative to the individual wordlist.

Furthermore, Rogers *et al*. (2011) found that measures of optimism derived from *Diction* had a positive correlation of 0.55 with the other two dictionaries, while optimism scores from Henry (2008) had a 0.32 correlation with those derived from Loughran and McDonald (2011). Thus, the measures obtained using different dictionaries appear to be measuring tone with different level of noise. Given the lack of any empirical evidence confirming the superior validity of any of the wordlists and the possibility that results are sensitive to the wordlist used, the choice of the appropriate wordlist becomes more challenging. Additionally, given

the importance of using context-specific wordlists to generate more valid measures of tone and other linguistic features of accounting documents (as highlighted in Loughran and McDonald, 2011), three wordlists were initially selected which were derived either partially (*Diction* wordlists) or fully (Henry 2008, Loughran and McDonald 2011) from business or financial documents. Subsequently, sensitivity analysis was conducted to ascertain the appropriateness of these lists for this study.

Diction is a Windows-based computer program that searches a passage of text for five main lexical features as well as thirty-five sub-features. It was created for the analysis of public discourse (political speeches) and the wordlists were derived from analysis of a sample of 20,000 texts from news coverage, advertisements, business reports including corporate financial reports, religious sermons and other documents (Hart, 2001). Mokoaleli-Mokotelli *et al*. (2009) provide an example of the use of *Diction* to study analysts' reports. Henry (2008) studied the tone of over 1,000 earnings press releases using a context-specific list of positive and negative words. This wordlist has also been used in other studies of accounting narratives such as in Rogers *et al*. (2011). Loughran and McDonald (2011) developed a finance related wordlist based on a large corpus of 50,000 10K documents which is now widely used within accounting (see for instance Twedt and Rees 2012). For comparison with previous studies, Table 28 presents example of studies of analysts' reports in which the various wordlists have been used.

Table 28: Wordlists and uses in analysts' research

Studies on analysts' reports	Wordlists
Fogarty and Rogers (2005), Mokoaleli-Mokotelli *et al*. (2009)	Diction
Kothari *et al*. (2009), Twedt and Rees (2012) and De Franco *et al*. (2015)	GI
None	LIWC
Twedt and Rees (2012)	LM
None	Henry

Sensitivity analysis

The aim of the sensitivity analysis is to test the validity of the measures of tone derived from each wordlist. The concept of validity and types of validity has been previously discussed in Chapter 4. The interest here is to test the construct validity of the measure of tone, i.e. the extent to which these wordlists accurately capture the tone in analysts' reports. The analysis was conducted using a random sample of reports obtained from the main sample used for the analysis. The software Nvivo[42] was used for this analysis. Unlike Diction, Nvivo enables the context of word use to be examined. The wordlists from Diction, Henry (2008) and Loughran and McDonald (2011) were used to create a text-search query and a node in Nvivo for positive and negative categories.

The sensitivity analysis involved two separate tests. The first involved checks for type I error or false hits. This is error due to falsely classifying a non-tonal word into a positive or negative category. This test involved checking the context (i.e. sentences or paragraphs) in which coded words are situated in the reports. A false hit is identified as one in which a word is classified as either negative or positive but in the context of its use in analysts' reports, it has a different meaning. This also includes the case of homonyms (i.e. words which have the same spelling with a dictionary term but have different meanings). The following extract illustrates a false hit:

> "Assets **under** custody and administration increased a bit to $22.8 trillion"
>
> *(Source: Credit Suisse report on State Street (STT), July 2011)*

After running the text search query to search for words which are contained in the wordlists, Nvivo highlights these words as in the extract above. The word 'under' is part of the negative wordlist from Henry (2008). However, in this setting, 'under' does not have a negative connotation. Assets under custody or assets under administration are common terms in fund management industry and do not

[42] See Chapter 4 for discussion of Nvivo and its functionalities.

have any positive or negative meaning in themselves. Thus, ‘under’ is considered a false hit in this setting.

Table 29 shows the number of words per report which were found to be false hits. In these cases, the words do not capture the construct of interest but are included in the measure of tone, limiting the validity of the measure. Appendix E shows further examples of coded and misclassified words for several reports.

Table 29: Misclassified words

Panel A: Negative

Reports	Dict_Neg			H_Neg			LM_Neg		
	Code	Miss	% Miss	Code	Miss	% Miss	Code	Miss	% Miss
1 : AMAT_MorningStar	33	17	52%	66	6	9%	66	1	2%
2 : AMAT_Oppenheimer	15	5	33%	9	1	11%	9	1	11%
3 : GOOG_CSuisse	22	18	82%	10	3	30%	10	3	30%
4 : GOOG_Indigo	11	3	27%	14	0	0%	19	0	0%
5 : SLB_Jeffries	16	8	50%	33	8	24%	38	10	26%
6 : SLB_WallSS	1	1	100 %	6	0	0%	8	2	25%
7 : STT_CSuisse	11	1	9%	15	6	40%	21	6	29%
8 : STT _MorningStar	36	3	8%	43	15	35%	47	6	13%
9 : WMT_WallSS	11	5	45%	19	3	16%	20	1	5%
10 : WMT_Wells	23	10	43%	34	0	0%	20	1	5%
Overall Total	**179**	**71**	**40%**	**249**	**42**	**17%**	**258**	**31**	**12%**
Avg % miss across reports			**45%**			**17%**			**15%**

Panel B: Positive

Reports	Dict_Pos			H_Pos			LM_Pos		
	Code	Miss	% Miss	Code	Miss	% Miss	Code	Miss	% Miss
1 : AMAT_MorningStar	58	18	31%	84	15	18%	74	4	5%
2 : AMAT_Oppenheimer	6	4	67%	9	4	44%	13	0	0%
3 : GOOG_CSuisse	49	3	6%	101	6	6%	35	0	0%
4 : GOOG_Indigo	36	4	11%	59	8	14%	27	2	7%
5 : SLB_Jeffries	59	11	19%	155	6	4%	90	1	1%
6 : SLB_WallS	13	2	15%	36	2	6%	22	1	5%
7 : STT_CSuisse	38	15	39%	75	17	23%	49	0	0%
8 : STT _MorningStar	39	14	36%	83	30	36%	28	2	7%
9 : WMT_WallSS	18	7	39%	24	4	17%	14	0	0%
10 : WMT_Wells	20	3	15%	53	2	4%	19	0	0%
Overall Total	**336**	**81**	**24%**	**679**	**94**	**14%**	**371**	**10**	**3%**
Avg % miss across reports			**28%**			**17%**			**3%**

Note: Panel A contains the results of the sensitivity analysis for the negative category from the Diction, H (Henry 2008) and LM (Loughran and McDonald, 2011) wordlists while panel B shows the results for the positive category. For each list, there are three columns showing the number of words coded to the positive or negative category (Code), the number of the coded words which are false hits (Miss) and the percentage of false hits over coded words (% Miss). Each report contains the ticker of the firm followed by an underscore and then the analyst-firm which produced the report. The overall total shows the total coded words, total misclassified words and the percentage of total misclassified words over total coded words (e.g. 71/179*100 = 40% for Dict_Neg). The average percentage of misclassified words across report is the mean of the %Miss column and measures the mean percentage of misclassified words across the reports.

Several salient observations can be made from this first set of tests. First, as shown in Table 29, coding using the Diction wordlist results in substantially more false hits than the other two lists (i.e. H and LM). About 40% of the total coded words are misclassified for the negative category and about 24% for the positive category compared to 17% and 14% for the H lists and 12% and 3% for the LM list. This supports the claim that context-specific wordlists provide more valid measures of tone compared to the generic lists in-built in most software programmes. Second, inspection of misclassified words (see Appendix E) reveals that several of the misclassified words have no negative or positive connotation in a financial context, e.g. "gross" and "fair" (as in fair value). Moreover, these words frequently occur in financial documents such as analysts' reports which poses a significant threat to the validity of the measures of tone if Diction were to be used.

The second set of sensitivity tests examined a sample of reports for cases of type II error, i.e. omission of relevant tonal words not contained in the wordlists. From Table 29 it can be noted that, apart from having fewer percentages of false hits, both the H and LM lists contain a higher number of coded words. This suggests that tonal words contained in analysts' reports are better represented in the H and LM lists compared to Diction. Thus, both of the specialist lists limits both type I and type II error compared to the Diction wordlist. Despite containing fewer words[43], the H list captures a similar number of tonal words and in some cases captures more tonal words than the LM list. For instance, the total number of coded words for positive category based on the H list is 679, as opposed to 371 for the LM list in Table 29. This indicates that the use of only LM list would have reduced the total number of words measuring positivity to 371, increasing type II error. Moreover, the H list seem to capture several tonal words typically found in analysts' reports, such as "beat", which is used to indicate whether firms' reported earnings beat managerial and/or consensus forecasts or prior expectations. Perhaps this is due to the nature of the documents from which the list is derived (i.e. earnings press releases) as they have some similarity with analysts' reports. Although authorship differs, both types of document often

[43] The H list contains 104 positive words and 85 negative words while LM lists contains 354 positive words and 2,329 words.

contain discussion and explanations of performance in comparison with benchmarks, e.g. consensus forecasts. Table 30 presents some examples of positive (Pos) and negative (Neg) words common to the LM and H lists (indicated by LM_H) and words often used in analysts' reports that are contained in the H list only (indicated by H_Only).

Table 30: Comparing Loughran and McDonald and Henry wordlists

LM_H (Pos)	H_only (Pos)	LM_H (Neg)	H_Only (Neg)
accomplish	beat	weak	below
achieve	above	challenge	down
enjoy	exceed	decline	drop
excellent	grew	risk	shrink
improve	solid	unfavourable	decrease
positive	up	difficult	fall

Words such as "above", "exceed", "up", "below", "down", "drop" are often used in contexts such as: "above consensus", "exceed management forecast", "up from prior year", "below our expectations", "down from last quarter", etc. All these are examples of performance comparisons commonly found in earnings press releases and analysts' reports. Hence, the H list appears to be more pertinent as these terms are included.

Further, the H and LM wordlist was compared. As detailed in Appendix E, coded tonal words vary across the different wordlists, indicating that both the LM and the H wordlists each provide unique words that capture tone. There are about 58 common words in the positive H and LM lists and 48 common words for the negative lists, which amount to 55% of the words in the positive H list and 56% for the negative lists. For the LM list, the overlapping words amount to 16% of the positive list and 2% of the negative list. This implies that the majority of words contained in the LM list are not included in the H list. Hence, to minimise type II error, a combined H and LM list is used to measure tone in this study and is given by:

$$Tone = \frac{(Positive\ words - Negative\ words)}{Total\ words};$$

where positive and negative words are based on a combined dictionary of Loughran and McDonald (2011) and Henry (2008). However, given that the H list is less established in the literature, the downside of this combined list is that

comparability with extant research is limited. This is mitigated by the use of the LM list as a robustness check.

Overall, the sensitivity analysis reported here indicates the importance of the contexts from which wordlists are derived. While the use of lists obtained from a business or financial contexts such as the LM list may be helpful in reducing misclassification error (type 1), it may not necessarily provide a significant improvement in capturing tonal words if the context of its use is different from the context (documents) in which it was derived from. Therefore, a criterion for selecting among the various lists for future researchers might be the similarity between the documents from which wordlists are derived and those for which the list is being used. For instance, a wordlist derived from a large sample of analysts' reports may contain additional words not captured by any of the existing wordlists as they originate from other corporate documents. A reading of the ten reports used for the sensitivity analysis discussed above revealed that all three wordlists failed to capture extra words such as "accelerate", "ahead", "upside", "surge", "raising", maximise", "robust", "rosy", "buoyed", "accretive", "well", "buy", "bullish", "soften", "bottomed", "reduce", "cuts", "dilutive", "pressure", "short", "red flag", "lowering", etc. These are words which were highlighted as either positive or negative while reading the reports. While some of these are unambiguously positive or negative such as "upside" others are not directional and depend on the context of use (e.g., "reduce").

Despite these sensitivity checks and the selection of a combined list for the analysis, the use of wordlists still presents several limitations. As previously mentioned, a key limitation of the use of a wordlist is that several words classified as either negative or positive are dependent on the context of their use and this is not captured in any case. For instance, the term "lower" is classified as a negative word in the chosen wordlists. The extracts below shows that in some contexts "lower" has a negative connotation e.g. "lower revenue", while in other context, it has a positive connotation such as "lower costs".

> "Applied has been streamlining operations to **lower** its cost structure."
>
> (AMAT_Morningstar)

> "Russia revenue was also **lower** with the onset of the winter slowdown."
>
> (SLB_Jeffries)

Thus, in one context, the word is misclassified as negative when it is indicating a positive outcome, while in the other it is correctly classified. The selection of the H and LM lists (or a composite list) does not and indeed cannot eliminate all misclassification errors due to homonyms or words which derive their meaning in context. For the analysts' reports, the context also extends to the subject being discussed. For instance, negative references to competitors' operations when analysing a company might be good news for the reporting company and should therefore be of a positive tone as opposed to a negative tone.

For instance, in the following extract, the chosen wordlists capture the word "less" as having a negative connotation. However, in context, it is referring to less diversified competitors from which the reporting company (State Street) is able to gain market share.

> "Management remains very bullish on the firm's business pipeline, which confirms our thesis that State Street is well-positioned to steal market share from **less** diversified local players."
>
> (STT_Morningstar).

Beyond the reporting company, analysts' reports contain references to other entities such as customers, suppliers, competitors, and the government and a negative or positive event for any of these may impact the operations of the reporting company in the opposite direction which might result in misclassification. Consequently, a more complex algorithm which captures contexts of word usage will further improve textual analysis. For the present study, the only solution has been to select more topic-related wordlists and measure the extent of misclassification which has been extensively discussed in the preceding paragraphs. As detailed in Table 29, the misclassification errors are not severe for the LM and H lists and the percentage agreement compares well with acceptable level of inter-rater reliability for manual content analysis as discussed in Chapter 4.

9.3.1.3 Data collection process

Reports uploaded onto Nvivo as describes in Chapter 5 were coded. However, unlike for the first research objective explained in that chapter, coding of words into positive or negative category was automatic in this case. The text query function in Nvivo was used. This involved creating a text query with the wordlists. Nvivo automatically runs the query and search the uploaded reports for the words contained in the query. All occurrences of the words are automatically coded to the positive or negative node category as specified in the query. Following coding, the number of positive words and negative words were extracted from the reports onto Microsoft Excel and the measure of tone computed as described above.

9.4 Reading ease

The second variable of interest for the second research objective is the readability of analysts' reports. As stated in Chapter 7, the differences in analysts' reports issued by IND and IB-analysts are examined to test for evidence of obfuscation in their writing. Courtis (2004, p.291) defines obfuscation as "a type of writing that obscures the intended message". Operationally, this is measured by a reading ease formula. However, the main challenge for research in this area is to identify a measure that is a reliable and valid proxy for reading ease. The methods used in prior accounting literature are briefly reviewed in Section 9.4.1 and the method of choice for this research is discussed in Section 9.4.2.

9.4.1 Measures of reading ease in prior research

Several studies within the field of accounting have examined the readability of different corporate reports. A highly-cited early review paper about these studies is Jones and Shoemaker (1994) who identified over 20 such studies. More recently, Merkl-Davies (2007) review other studies and identified three main methods used to study readability in prior literature:

9.4.1.1 Text-centred approach

In this approach, readability is assumed to be a function of the text and therefore, readers' backgrounds, education and interests are ignored. This takes the form of quantitative analysis in which a numerical score is computed as a measure of readability. Usually readability is judged based on the syntactic characteristics of the text, such as word length, sentence length and complexity of the words used. This is the basis for most readability measures such as the Flesch-Kincaid reading ease formula, Fog index, Lix, Dale and Chall etc. These are often used individually or aggregated to provide an average readability score (see for instance De Franco *et al.*, 2015).

9.4.1.2 User-centred approach

In this approach, reading ease is assumed to be a function of the reader rather than the features of the text read. Generally, the reader's understanding is examined which is in part contingent upon syntactical complexity. The Cloze index

provides an example of the user -centred approach to examining reading ease. It involves the deletion of every n^{th} word in a piece of text and requiring the reader to fill in the gap. Rather than reading ease, this approach is more likely to measure understandability which is a different but related concept (Jones, 1997).

9.4.1.3 Integrative approach

The final approach identified by Merkl-Davies (2007) combines the previous two approaches. A typical example is offered by the texture index introduced to the accounting literature by Sydserff and Weetman (1999).

9.4.2 Approach used in this research

In this study, a text-centred quantitative approach is used as readability of analysts' reports is measured by a reading ease score. This is consistent with the objective of this section of the thesis which is concerned with whether analysts bias their writing style to conceal bad news. Hence, a measure of readability based on the textual features of the text or the writing style of the author is desirable. It is judged that a user based approach is more concerned with characteristics of the users rather than the source and this is beyond the scope of this present study. Given the myriad of reading ease measures available and used in extant accounting literature, the next challenge is the selection of a formula which is a reliable and valid measure of syntactic complexity and hence, readability. The Fog index and the Flesch index are most frequently used in the accounting literature to measure corporate reports' readability (Merkl-Davies, 2007; Linsley and Lawrence, 2007). Both are based on sentence length and the length of the words contained in the sentences and are given by the following formulae:

Fog Index = (average number of words per sentence + percent of complex words)× 0.4

where complex words are measured as words with three syllables or more.

$$Flesch - Kincaid = (11.8 \times Syllables\ per\ word) + (0.39 \times words\ per\ sentence) - 15.59$$

The underlying assumption is that reading ease is a function of sentence and word length. These variables may not represent the construct of readability reliably but are relatively assumed to be associated with readability (Lewis et al., 1986). As discussed in Courtis (1998, p. 460) "It is believed that word length is related to speed of recognition, and that sentence length is related to memory span or words recalled."

Despite their popularity and the objectivity and reliability with which these measures are derived, they are still fraught with several weaknesses which threaten their validity as highlighted in previous studies. First, it is argued that these measures are too simplistic as they measure reading ease based on word and sentence length while ignoring other relevant factors such as style, syntax and cohesion and could provide misleading results (Courtis, 1998). Moreover, the measures of word difficulty are based on length of the word as measured by number of syllables rather than familiarity with the word. Loughran and McDonald (2014) highlights that these formulae misclassify and penalise several financial words as complex though easy to understand to investors. For example, they found that multisyllabic words such as 'telecommunications' which are penalised by the Flesch and Fog indices are actually not difficult to understand. Second, these measures were originally derived within the context of children writing and their appropriateness for evaluating technical accounting reports are questionable (Sydserff and Weetman, 1999). Third, they were developed over 50 years ago and thus fail to account for changes in language over time (Clatworthy and Jones, 2001).

These reasons have led to a search for alternative measures. Loughran and McDonald (2014) advocated the use of file size to measure readability of annual reports. They justified the use of file size over extant measures because "it is straightforward, is substantially less prone to measurement error, is easily replicated, is strongly correlated with alternative readability measures" (p.1644). However, as argued by Bonsall *et al.* (2015), there is no theoretical justification for the use of file size as a measure of readability. Moreover, given that obfuscation of bad news may also take the form of omission of disclosing relevant information, longer or larger (in terms of file size) documents may contain more information and thus be indicative of less concealment. Additionally, the length

and structure of analysts' reports in the sample varies with brokerage house and research firm such that the use of report length or size in this context may capture other variables not necessarily associated with readability.

Arguably a more valid measure of readability should measure features of text that are theoretically associated with the concept of readability. The US Securities and Exchange Commission encourages firms to use 'plain English' in all communications with investors. To provide some practical guidance on how to enhance the readability of financial disclosures, it issued a handbook in 1998 entitled "*A plain English handbook: How to create clear SEC disclosure documents*". Recommendations were based on a review of regulatory filings with the SEC and consultation with experts. The recommendations include issues such as less use of:

- Passive voice
- Hidden verbs
- Superfluous words
- Legal and financial jargon
- Defined terms
- Abstract words
- Lengthy sentences

They provide examples of each case as well as rational for their impact on readability. They argued that financial disclosures containing less of these features will be more readable. Thus, a valid readability measure should capture these multidimensional attributes. One such measure is the *Bog index*, which was developed professionally as part of the plain English editing software, StyleWriter.

Bog index

This is a measure of readability that assesses several features of text that "bogs" down the reader. These features include those conceptually related to readability as highlighted in SEC (1998) and presented above. The formula is given by: $Bog\ Index = Sentence\ Bog + Word\ Bog - PEP$,

where higher values imply less readable text. Each component of the formula is discussed briefly below:

Sentence Bog: measures the attributes of a text associated with sentence composition that limits readability, i.e. sentence length. It is given by:

$$Sentence\ Bog\ = \frac{(average\ sentence\ length)^2}{long\ sentence\ limit}$$

The default long sentence limit is 35 words per sentence. However, the software enables adjustment of this limit based on type of text and the intended audience.

Word Bog: measures attributes of a text associated with word difficulty, which limits the readability of a text. Unlike the Fog and the Flesch indices which measures word difficulty based on the number of syllables, the *Bog index* measures word difficulty based on familiarity and it also includes other aspects included in SEC (1998) and is given by the following formula:

$$Word\ Bog\ = \frac{(style\ problems + heavy\ words + abbrevitions + specialist\ words) \times 250}{number\ of\ words}$$

Style problems include a range of features highlighted in SEC (1998) including passive verbs, hidden verbs, wordy phrases etc. 'Heavy' words refer to difficult or complex words. Unlike the Fog and Flesch scores, in which Word difficulty is assumed to be based on its length, in the case, word difficulty is based on familiarity with the words. The Stylewriter software uses a graded dictionary of about 200,000 words. Each word is graded as difficult or easy based on familiarity/ease of recognition. Bonsall *et al*. (2015) tests the validity of Stylewriter' word grading system. They identify about 100 multisyllabic words, commonly found in financial documents very familiar to the average investor. They found that these common and highly used words are less penalised using the *Bog index* than the *Fog index*. Abbreviations are abbreviated words and acronyms and specialist words are technical terms.

Finally, the *Bog index* rewards features that make a text more readable. This is measured by the PEP component of the formula and is deducted from the sentence and word bog. PEP includes the use of proper names, interesting words and

conversational expressions such as use of personal pronouns and variation in sentence length. It is given by:

$$PEP = \frac{(names + interest\ words + conversational) \times 25}{number\ of\ words} + sentence\ variety$$

Where $Sentence\ variety = \frac{Standard\ deviation\ of\ sentence\ length \times 10}{average\ sentence\ length}$

Finally, the *Bog index* assumes that readability depends in part on the audience and the Stylewriter software permits the user to adjust the writing and audience. There are a variety of writing tasks and the audience could be 'in-house', 'public' or 'specialist'. A writing tailored to a specialist audience will impose less penalty on the use of specialist words compared to one directed to the public.

Given the features described above, the *Bog Index* provides a superior measure of readability and overcome common problems of the Fog and Flesch index. First, it captures features of text that are conceptually related to readability as highlighted in SEC (1998). Bonsall *et al.* (2015) provides a comparison of these features as contained in Table 31.

Table 31: Comparison of SEC (1998) and Bog Index' plain English features

SEC plain English problem	Stylewriter plain English components
Passive voice	Passive verbs
Hidden verbs	Hidden verbs
Superfluous words	Wordy phrases
Legal and financial jargon	Legal words and jargons
Defined terms	Specialist words
Abstract words	Abstract verbs
Unnecessary details	*Bog index* and number of words
Long sentences	Long sentences
Unreadable design and layout	N/A

Source: adapted from Bonsall *et al.* (2015)

Second, word difficulty is based on familiarity and thus more valid than simple word length as being used in the Fog and Flesch scores. Third, it measures both good (PEP) and bad (sentence and word bog) features of good writing and thus is more comprehensive than other measures of readability. Finally, it adjusts for

writing task and the audience of the document and overcomes the limitations of the Fog and Flesch index which were used for grading children text and do not offer any adjustment for documents such as financial disclosures. Although relatively novel, the *Bog index* and other features of the Stylewriter software have been used in academic studies within the accounting literature (see, for instance, Miller, 2010; Rennekamp, 2012 and Bonsall *et al*. 2015).

9.5 Chapter Summary

This chapter was aimed at describing the choices made with regards to the selection of an appropriate wordlist for measuring tone as well as the choice of most appropriate measure of readability.

First, the approaches used in extant literature to study linguistic features of accounting documents was reviewed and it was highlighted that this research is based on quantitative approaches such as word counts.

Second, research design choices relating to the selection of an appropriate word list were discussed. This involved a review of the wordlists currently used and their appropriateness for this research and test of their validity and limitations. The word lists by Henry (2008) and Loughran and McDonald (2011) were found to yield more valid measures of tone and were used for this research.

Third, the limitations of the existing measures of readability were discussed and a relatively new and professional readability measure (*Bog index*) was introduced as an alternative score used in this study. It was found to capture several features conceptually related to readability as contained in SEC's plain English handbook (SEC, 1998).

Chapter 10: Variation in the linguistic features of analysts' reports: Empirical results and analysis

10.1 Introduction

This chapter presents and discusses the results of the empirical analysis undertaken to address the second research objective of the thesis, which is concerned with variation in linguistic features (readability and tone) of analysts' reports. The aim of the analysis is to investigate the factors that explain the variations in these linguistic features across analysts' reports. The first step in the data analysis process was the production of detailed descriptive statistics for the dependent (measures of tone and readability), which are presented and discussed in Section 10.2. This is followed by univariate analysis in section 10.3. Test of hypotheses include paired sample t-tests, used to examine differences in linguistic features between IB and IND-analysts' reports. This is corroborated with conditional logistic regression. A multivariate analysis is also used to test association between linguistic features and company characteristics. The results of the hypotheses tested are presented and discussed in Section 10.4 and Sensitivity analysis is discussed in Section 10.5.

10.2 Descriptive statistics

This section presents and discusses descriptive statistics for the variables used in this study. This section focuses mostly on the dependent variables as the statistics for the independent variables consist of company characteristics which have been previously discussed in Chapter 6 of this thesis and is not restated here.

The dependent variables measure the linguistic features of analysts' reports (i.e. the tone of the reports and the syntactic complexity or readability of the reports). The descriptive statistics for these variables are presented in Table 32.

Table 32: Descriptive statistics of measures of tone and readability

Variables	N	Mean	Std. Dev	Min	Median	Max
Bog_IND	144	57.11	14.59	32.00	55.00	108.00
FleschK_IND	144	15.45	1.76	11.30	15.55	19.00
Fog_IND	144	14.16	2.17	7.50	14.35	19.00
Tone_IND	144	2.60	1.81	-1.90	2.45	8.04
Tone_LM_IND	144	0.60	1.17	-2.48	0.50	4.30
Bog_IB	144	69.40	16.13	21.00	70.00	107.00
FleschK_IB	144	15.94	2.27	1.30	16.10	19.00
Fog_IB	144	12.98	2.02	8.30	12.90	18.10
Tone_IB	144	2.41	1.85	-1.92	2.23	7.92
Tone_LM_IB	144	0.41	1.09	-3.00	0.47	3.39

This table shows summary statistics for the linguistic features of the sampled reports. The suffix _IB and _IND represents data from reports issued by investment banking and brokerage analysts and analysts employed at independent research firms, respectively. 'Bog' is the *Bog index* which is used to measure the readability of the reports in this study. 'Tone' measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). This is calculated as the number of positive words less the number of negative words divided by the total word count and multiplied by 100. 'Tone_LM' is calculated similarly but is based only on the Loughran and McDonald (2011) word list. 'Fog' is the Fog index and FleschK is the Flesch Kincaid index. The highlighted rows are the main variable of interest used in this study. Others have been included to only enable comparison with other studies.

As previously highlighted in Chapter 9, the *Bog index* is increasing in readability such that the higher the *Bog index*, the more complex or less readable the reports. The *Bog index* score for IB-reports ranges from 21 to 107 which is comparable to the IND-reports, which range from 32 to 108. Despite the similarity in range, reports issued by IB-analysts have a markedly higher mean (median) of 69.4 (70)

compared to a mean (median) of 57.11 (55) for reports issued by IND-analysts. To enable comparison with other studies, the descriptive statistics for the *Fog* and *Flesch Kincaid* scores are also reported. Using the *Fog* index and *Flesch - Kincaid* scores as measures of readability, the mean (median) are 12.98 (12.9) and 15.94 (16.1) for IB-reports and 14.16 (14.35) and 15.45 (15.55) for IND-reports respectively. Using the Fog index for a sample of 2,057 and 356,463 reports, Twedt and Rees (2012) and De Franco *et al*. (2015) report a mean of 16.96 and 18.7, respectively. Based on the Fog index, the results in the present study are lower compared to these studies. First, unlike Twedt and Rees (2012) who examined initiating coverage reports, the present study only examined results reports. The differences in the *Fog* index may be explained by the differences in scope, length and content of both types of reports. Initiating coverage reports are usually lengthier and to the extent that length of report is indicative of complexity (Li, 2008), the higher score found in Twedt and Rees (2012) may be expected. Second, De Franco *et al*. (2015) did not distinguish between types of reports while the present study focuses on results reports only, as well as distinguishes between IB and IND-analysts' reports. Consequently, the differences may be partly explained by these differences.

For the tone variable, the mean (median) for IB-reports is 2.41 (2.23) and 2.60 (2.45) for IND-reports using the combined dictionary of Loughran and McDonald (2011) and Henry (2008). The mean value is positive for both IB and IND sub-sample and suggests that analysts' in general tend to issue reports with more positive than negative words. However, the mean score is higher for IND-reports in the sample. The results persist when tone is measured using only the Loughran and McDonald (2011) wordlist (*Tone_LM_IB* and *Tone_LM_IND*). While the mean score remain positive, the values are lower suggesting the Loughran and McDonald (2011) wordlist only captures a lesser proportion of tonal words compared to the combined list as previously discussed in Chapter 9.

In general, the optimistic tone reflected in the sampled reports is consistent with prior studies such as Twedt and Rees (2012), who report a mean value for tone of

4.16 using the GI software[44]. This is also consistent with prior studies which document analysts' tendency to optimism through the issuance of more positive recommendations or earnings forecasts as reviewed in Chapter 7. The results further suggest that analysts' tendency to bias may not be limited to the summary output of their research process such as recommendations and earnings forecasts (though this has been the focus of prior research), but may also be reflected in the narratives that accompany these measures.

10.3 Univariate analysis

In this section, univariate analysis consisting of bivariate correlation between the linguistic features and company characteristics are reported. Table 33 presents the Pearson correlation coefficients.

[44] Their measure of tone is defined in the same way as used in this study, i.e. (positive word count – negative word count)/ total word count *100, although a different word list is used.

Table 33: Pearson correlation matrix

	ΔP_EPS	Size	Leverage	Risk	Growth	Intan	MTB	BOG_IB	Tone_IB	Bog_IND	Tone_IND
ΔP_EPS	1										
Size	-0.03	1									
Leverage	-0.02	-0.15*	1								
Risk	0.06	-0.132	-0.17**	1							
Growth	-0.09	0.224***	-0.22**	-0.06	1						
Intan	-0.15*	0.018	0.13	-0.23***	0.12	1					
MTB	-0.05	-0.059	0.26***	0.01	0.06	-0.04	1				
Bog_IB	-0.01	-0.02	-0.11	-0.06	0.08	0.12	0.04	1			
Tone_IB	0.07	0.004	-0.07	-0.17**	0.09	0.08	-0.06	-0.20**	1		
Bog_IND	-0.07	0.071	-0.16*	0.022	0.38***	0.14	-0.04	0.29***	-0.12	1	
Tone_IND	0.25***	0.024	0.05	-0.16*	-0.07	0.06	0.05	-0.05	0.39***	-0.11	1

This table shows the Pearson correlation coefficients for the linguistic features and company characteristics. The suffix _IB and _IND represents data from reports issued by investment banking and brokerage analysts and analysts employed at independent research firms. *Bog* is the *Bog index* which is used to measure the readability of the reports in this study. *Tone* measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). This is calculated as the number of positive words less the number of negative words divided by the total word count and multiplied by 100. *ΔP_EPS* is the percentage change in EPS from prior year. *Size* is the natural logarithm of market value for the fiscal year end 2010. *Leverage* is the long term debt to total asset ratio. *Risk* is the standard deviation of EBIT over five years (2006-2010) scaled by average assets over the time period. *Growth* is the cumulative annual growth rate in sales per share over the previous five years from 2006-2010. *Intan* is the proportion of intangible assets over total assets. *MTB* is the ratio of market value of equity to book value of equity for the fiscal year end 2010. All variables were obtained from Computstat. ***/**/* means significance at 1%, 5% and 10% respectively.

The results in Table 33 reveal some significant correlations between the independent variables which measure several company characteristics. These have been previously highlighted and discussed in Chapter 6 of the thesis and are not reviewed here. Perhaps, most importantly is that the correlation between the variables are small in magnitude (being all below 0.5), suggesting the associations are not strong enough to pose a multi-collinearity threat in the multivariate analysis reported in subsequent sessions. Kennedy (2008, p. 196) suggest that a value of 0.8 and above is useful indication of collinearity between variables.

Table 33 also show that most company-specific variables are not significantly associated with measures of tone and readability of analysts' reports. The main variable of interest for test of hypotheses is the measure of financial performance, which is proxied by *ΔP_EPS*. Only the tone of IND-reports are correlated with financial performance. The association is positive (with a coefficient of 0.25) and is significant at the 1% level. This provides some support for hypothesis H_{E2} and suggests that tone of analysts' reports is based in part on changes in financial performance of the firm being covered by the analysts with more positive reports being issued for firms with positive changes in EPS. *Risk* is negatively associated with measures of tone at the 5% level for *Tone_IB* and at the 10% level for *Tone_IND*, Indicating that analysts' are more optimistic of firms with less variable financial performance. Additionally, *Growth* is positively associated with the *Bog index* for IND-reports, with a correlation coefficient of 0.38, which is significant at the 1% level. This suggests that IND-reports for high growth firms are more complex (i.e. less readable as higher score for the *Bog index* indicate more complex or less readable reports). Overall only few company attributes are significantly correlated with the linguistic features and may suggest that company attributes are not the main factors which explain variation in the tone and readability of analysts' reports.

The tone and readability of reports issued by IB and IND-analysts are correlated. First, the tone of IB-reports is negatively associated with the *Bog index*, indicating that reports with lower tone are less readable (i.e. have higher *Bog index*). The coefficient is -0.2 and is significant at the 5% level. This provides support for hypotheses H_{G1} and indicates that when analysts' do not have good news to report, their reports are less readable. This pattern is not so for IND-reports. Second,

Tone_IB is correlated with Tone_IND at the 1% level with a coefficient of 0.39 and Bog_IB is positively associated with Bog_IND at the 1% level with a coefficient of 0.29.The positive sign of the coefficients suggest that there is a common influence on the sources of their variation, however, the magnitude of the coefficients is small, being lower than 0.5, indicating that there are still differences not explainable by common factors.

10.4 Hypotheses testing

This section presents the results of test of the pre-specified hypotheses. These hypotheses test the differences in the use of linguistic features between IB and IND-analysts and their association with company characteristics such as financial performance.

10.4.1 Hypotheses $H_{1 \text{ and } 2\,A}$: Differences in tone and readability of IB and IND-analysts' reports.

A 1-1 matched sample research design is used to test for differences in linguistic features between IND and IB-analysts' reports. Given the differences in the incentives to issue optimistic research, H_{1A} and H_{2A} test whether IB-analysts' reports use more positive words and are less readable compared with IND-analysts' reports (for poor performing firms). A paired sample t-test of differences and conditional logistic regression are used to test these hypotheses. Results are presented in Table 34 and 35.

Table 34: Paired sample test of difference in tone

Panel A: Difference in mean values			
Variable	**IB**	**IND**	**Diff**
Tone	2.41	2.60	-0.18
Panel B: Difference in median values			
Variable	**IB**	**IND**	**Diff**
Tone	2.23	2.45	-0.22

This table shows the results of a paired sample test of difference in mean and median scores of the measures of tone between IB and IND-reports. Panel A reports test of difference in the mean values using the paired sample t-test, while Panel B reports test of difference in median values using the non-parametric equivalent of the t-test, the Wilcoxon signed rank test. The total number of observations is 288 (144 IB-analysts' reports and 144 IND-analysts' reports). *Tone* measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). ***/**/* means significance at 1%, 5% and 10% respectively.

Table 35: Paired sample difference in linguistic features - conditional logit

IB	Coefficient	Odds Ratio	Z	P-value
Bog	0.07	1.07	5.72	0.00
Tone	-0.003	1.00	-0.03	0.98
POS	-0.14	0.87	-0.45	0.65
N	288			
Wald X^2	33.98			
Significance	0.00			
Pseudo R^2	0.26			

This table shows the results of using the conditional logistic regression model to test for differences in the linguistic features of IB and IND-analysts' reports. The dependent variable is the logit of the odds of a report being issued by an IB-analyst. The variable *IB* takes the value of 1 for reports issued by an IB-analyst and 0 for reports issued by an IND-analyst. *Bog* is the *Bog index* which is used to measure the readability of the reports in this study. Higher values indicate more syntactic complexity and thus less readable reports. *Tone* measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). POS is a dummy variable which takes the value of 1 for a positive recommendation and 0 for a neutral or negative recommendation and used as a control variable in this setting. The odds ratio is the exponentiated coefficient. Z is the z-value based on the Wald test of significance of each of the coefficients. The significance of the overall model is tested by the Wald X^2, which test whether the coefficients are jointly equal to zero. ***/**/* means significance at 1%, 5% and 10% respectively.

Holding the company characteristics constant by using a matched sample of IB and IND-analysts' reports for the same companies, Table 34 show that the difference in the tone of IB and IND-report is not statistically significant. Contrary to expectation of a higher tone for reports issued by IB analysts as stated in H_{1A}, the results show that IND-analysts issue reports which are more optimistic with a mean value for the tone variable of 2.6o *c.f.* 2.41 for IB-analysts. The results are maintained for non-parametric tests of differences in median as confirmed in Panel B of Table 34. Further, conditional logistic regression was performed to ascertain the effect of tone on the likelihood of a report being issued by an IB-analysts, controlling for report readability and the recommendation level. The results are confirmed by the conditional logistic model presented in Table 35 as the coefficient on tone is not statistically significant. Although, H_{1A} is not supported as differences in tone are not significant, the mean value is positive across IB and IND-reports and is indicative of general tendency towards optimism by both analysts. This tendency is also reflected in the higher proportion of positive recommendations compared with negative recommendations in the sample. This result is supportive of the findings in several studies which document that IB-analysts issue earnings' forecasts which are more accurate and less optimistic compared with non-IB analysts (e.g. Jacob *et al.* 2008). Jacob *et al.*

(2008) argued that IB-firms employ higher quality analysts which may be concerned about reputational damage of producing optimistic research to curry favour with managers. Moreover, IND-analysts, whose research reports are paid for and sponsored by corporate issuers, may be more incentivised to curry favour with corporate managers given their dependency on the fees (Kirk, 2011). This result extends and complement these studies by providing additional evidence with regards to the tone of analysts' reports. It also further show that although IB-analysts' reports are no more optimistic than IND-reports, they tend to make their reports less readable. The findings lend some support for hypotheses $H_{2A,}$ although the result presented are for the unconditional test of difference in readability. Hence, Table 36 and 37 present the results of paired sample differences in readability for poor performing firms (i.e. firms for which changes in financial performance, as measured by percentage change in EPS from prior year, is negative), which is aimed at testing hypotheses $H_{2A.}$

Table 36: Difference in readability for poor performing firms

Panel A: Difference in mean values			
Variable	**IB**	**IND**	**Diff**
Bog	70.25	56.07	14.18***
Panel B: Difference in median values			
Variable	**IB**	**IND**	**Diff**
Bog	70.5	57	13.5**

This table shows the results of a paired sample test of difference in mean and median scores of the measures of tone and readability between IB and IND-reports for a sub-sample of companies for which changes in EPS from prior year is negative. Panel A reports test of difference in the mean values using the paired sample t-test, while Panel B reports test of difference in median values using the non-parametric equivalent of the t-test, the Wilcoxon signed rank test. The total number of observations is 56 (28 IB-analysts' reports and 28 IND-analysts' reports). *Bog* is the *Bog index* which is used to measure the readability of the reports in this study. Higher values indicate more syntactic complexity and thus less readable reports. ***/**/* means significance at 1%, 5% and 10% respectively.

Table 37: Paired sample difference in readability for poor performing firms -conditional logit

	Coefficient	Odds Ratio	Z	P-value
BOG	0.07	1.08	2.17	0.03
Tone	-0.12	0.88	-0.40	0.69
POS	-1.29	0.27	-1.36	0.17
N	56.00			
X^2	8.39			
Prob > X^2	0.04			
Pseudo R^2	0.37			

This table shows the results of using the conditional logistic regression model to test for differences in readability of IB and IND-analysts' reports. The dependent variable is the logit of the odds of a report being issued by an IB analyst. The variable *IB* takes the value of 1 for reports issued by an IB analyst and 0 for reports issued by an IND analyst. *Bog* is the *Bog index* which is used to measure the readability of the reports in this study. Higher values indicate more syntactic complexity and thus less readable reports. *Tone* measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008) and is used as a control variable in this case. POS is a dummy variable which takes the value of 1 for a positive recommendation and 0 for a neutral or negative recommendation and used as a control variable in this setting. The odds ratio is the exponentiated coefficient. Z is the z value based on the Wald test of significance of each of the coefficients. The significance of the overall model is tested by the Wald X^2, which test whether the coefficients are jointly equal to zero. ***/**/* means significance at 1%, 5% and 10% respectively.

The results reveal that for firms with negative changes in financial performance, IB-analysts issue less readable reports. The paired difference is positive and significant for the *Bog index*. Similarly, in the conditional logit model, the *Bog index* has a significant coefficient. Thus, hypotheses H_{2A} is supported by the results.

10.4.2 Hypotheses $H_{1 \text{ and } 2B}$: Differences in association between tone, readability of IB and IND-analysts' reports and financial performance

Hypotheses H_{2A} and H_{2B} tests whether tone and readability of IB-reports are differently associated with financial performance compared with IND reports. These hypotheses are tested using multivariate analysis as modelled in section 8.5.1 of the thesis. The principal independent variable of interest is financial performance as measures by percentage change in EPS and other variables are included for control. The results are presented in Tables 38 and 39.

Table 36: OLS regression results of tone and company characteristics

	Tone_IB			Tone_IND			Tone_All	
	Coeff.	**T-values**		**Coeff.**	**T-values**		**Coeff.**	**T-values**
Intercept	3.10	1.46		0.63	0.44		2.04	1.25
P_EPS	0.08	3.19**		0.25	3.44***		0.24	7.03***
Size	-0.12	-0.58		0.14	1.01		-0.01	-0.07
Leverage	-1.26	-1.83		0.32	0.21		-0.46	-0.54
Risk	-9.46	-3.59***		-7.25	-1.31		-8.51	-3.09***
Growth	2.06	1.32		-0.68	-0.51		0.68	0.77
Intan	1.18	1.81		1.12	1.35		1.13	2.01*
MTB	0.00	-0.57		0.01	1.17		0.00	0.7
POS	0.48	1.66		0.48	1.48		0.54	1.84*
LTC	0.81	2.3**		0.29	0.99		0.57	2.1*
IB							-0.07	-0.24
IB*P_EPS							-0.16	-3.79***
N		132			128			264
F		3.91			2.82			14.45
P-value		0.00			0.00			0.00
R^2		0.05			0.07			0.08

This table shows the results of multivariate analysis of tone regressed on company characteristics which is aimed at testing the association between financial performance (measured by P_EPS) and tone. Three models were estimated. The first model is for the IB-analyst sub-sample and the second model is for the IND-analysts sub-sample and the third model is for the entire sample containing both IB and IND-analysts' reports. The entire sample is 288 reports and 144 for IB and IND analysts sub-samples. The smaller sample size reported here is a result of missing variables. The dependent variables in all models is *Tone*, which measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). The dependent variables are: P_EPS is the percentage change in EPS from prior year. Size is the natural logarithm of market value for the fiscal year end. Leverage is the long term debt to total asset ratio. Risk is

the standard deviation of EBIT over five years scaled by average assets over the time period. Growth is the cumulative annual growth rate in sales per share over the previous five years. Intangibles is the proportion of intangible assets over total assets. MTB is the ratio of market value of equity to book value of equity for the fiscal year end. POS is dummy variable which takes the value of 1 for reports with a positive recommendation and 0 otherwise. Companies are classified into two industry groups, i.e.hi-tech (HTC) and low-tech (LTC). Classification of firms into HTC and LTC is based on the classification scheme used in Francis and Schipper (1999). LTC is dummy variable which takes the value of 1 for reports for firms in the low-tech industry and 0 otherwise. IB is dummy variable which takes the value of 1 for reports by investment banking analysts and 0 otherwise. IB*P_EPS is an interaction term between IB and P_EPS. ***/**/* means significance at 1%, 5% and 10% respectively.

The argument at the basis of hypothesis H_{2A} is that in a situation in which the tone of analysts' report is used strategically to ingratiate corporate managers (as previous studies have documented for recommendations and earnings' forecasts), the tone will be less associated with changes in financial performance. Thus the tone of IB-analysts should be less associated with financial performance compared to IND-analysts. The results in Table 38 reveal that across all models, the coefficient on P_EPS is positive and significant at the 5% level or better. For the IB-model, a percentage increase in EPS from prior year increases the optimism of the reports and the tone variable increases by 0.08 with a *p-value* <0.05. For the IND-model, a percentage increase in EPS from prior year increases the optimism of IND-reports by 0.25 and is significant at the 1% level. Though both coefficients are significant and positive, the magnitude is higher for the IND-model. To test whether these difference in coefficient is significant, the regression was estimated for the combined sample with an interaction variable (IB*P_EPS) which test whether the difference in coefficient is higher or lower for the IB only sub-sample compared with the IND sub-sample. The coefficient on the interaction variable is negative and significant at the 1% level. It suggest that although the tone of analysts' report in general is positively associated with the changes in financial performance (with a positive and significant coefficient on P_EPS), the effect is lesser for IB-analysts compared with IND-analysts, providing support for hypothesis H_{2A}. These results imply that IB-analysts, who are hypothesised to be more predisposed to ingratiating corporate managers write reports which are less associated with company fundamentals such as financial performance. This provides support for optimistic bias among IB-analysts.

Hypotheses H_{2B} test whether the readability of analysts' reports vary with changes in financial performance. The premise is based on findings that corporate reports are less readable for low level of financial performance (Li, 2008). Thus, in a

situation of negative financial performance, analysts' predisposed to ingratiating managers will produce less readable reports. Table 39 presents the results of multivariate analysis.

Table 37: OLS regression results of readability and company characteristics

	Bog_IB			Bog_IND			Bog_All	
	Coeff.	**T-values**		**Coeff.**	**T-values**		**Coeff.**	**T-values**
Intercept	82.84	4.67***		71.52	4.92***		69.21	7.03***
P_EPS	0.13	0.26		-0.14	-0.25		-0.18	-0.49
Size	-0.81	-0.44		-0.98	-0.65		-0.64	-0.54
Leverage	-18.21	-3.46***		-5.06	-0.54		-11.41	-1.75
Risk	-33.26	-0.92		8.09	0.26		-11.30	-0.58
Growth	1.09	0.09		40.65	3.36***		20.87	2.6**
Intan	8.14	0.97		0.33	0.06		4.59	0.92
MTB	0.07	6.06***		-0.05	-0.96		0.01	0.36
POS	-1.30	-0.47		-0.29	-0.12		-1.61	-0.77
LTC	-3.74	-1.41		-9.37	-3.57***		-6.80	-4.06***
IB							12.24	3.35***
IB*P_EPS							0.37	0.57
N		132			128			264
F		1.19			4.51			62.13
P-value		0.00			0.00			0.00
R^2		-0.02			0.17			0.20

This table shows the results of multivariate analysis of Bog index regressed on company characteristics which is aimed at testing the association between financial performance (measured by P_EPS) and readability. Three models were estimated. The first model is for the IB analyst sub-sample and the second model is for the IND analysts sub-sample and the third model is for the entire sample containing both IB and IND analysts' reports. The entire sample is 288 reports and 144 for IB and IND analysts sub-samples. The smaller sample size reported here is a result of missing variables. The dependent variable in all models is *Bog* which measures the readability of the reports using the *Bog index*. The dependent variables are: P_EPS is the percentage change in EPS from prior year. Size is the natural logarithm of market value for the fiscal year end. Leverage is the long term debt to total asset ratio. Risk is the standard deviation of EBIT over five years scaled by average assets over the time period. Growth is the cumulative annual growth rate in sales per share over the previous five years. Intangibles is the proportion of intangible assets over total assets. MTB is the ratio of market value of equity to book value of equity for the fiscal year end. POS is dummy variable which takes the value of 1 for reports with a positive recommendation and 0 otherwise. Firms are classified into two groups, i.e.hi-tech (HTC) and low-tech (LTC). Classification of firms into HTC and LTC is based on the classification scheme used in Francis and Schipper (1999). LTC is dummy variable which takes the value of 1 for reports for firms in the low-tech industry and 0 otherwise. IB is dummy variable which takes the value of 1 for reports by investment banking analysts and 0 otherwise. IB*P_EPS is an interaction term between IB and P_EPS. ***/**/* means significance at 1%, 5% and 10% respectively.

The results reported in Table 39 show that readability is not significantly associated with financial performance (as measured by P_EPS) across all models. The interaction variable in the combined model (IB*EPS) is also not significant indicating that the association between the readability of the reports and changes in EPS is not significantly different across between IB and IND-analysts. Thus, hypotheses H_{2B} is not supported by the results. One possible explanation is that the financial performance is measured narrowly by the use of changes in EPS. Additional test was conducted in which the P_EPS variable was replaced with an aggregate score which comprised percentage changes in EPS, revenue, operating profit margin and return on assets. The financial performance coefficient remained insignificant.

10.4.3 Hypotheses H_3: Differences in tone and readability for optimistic recommendations

Hypotheses H_3 test for differences in tone and readability when the sample is limited to cases in which IB-analysts are relatively optimistic in their recommendations compared with IND-analysts and vice- versa. As argued in Chapter 8, by limiting the sample to cases of relative optimistic recommendation, H_3 provide a stronger test of whether analysts' strategically bias their reports. In the first instance, the analysis is limited to those firms for which IB-analysts issue positive recommendations which are not accompanied by a positive recommendation from the IND-analyst following the same company. Given the paired sample design, differences in company characteristics that may explain the differences in recommendation types are inherently controlled for. Consequently, a positive recommendation by an analyst which is not supported by another may suggest a case of (relative) optimistic bias. It is argued that if cases of relative optimism in recommendation are motivated by incentive to curry favour with managers, the accompanying narratives will show evidence of reporting bias (as argued by Malmendier and Shantikumar, 2014). Given IB-analysts' greater incentive to ingratiate (i.e. curry favour with managers), it is expected that these features will be more evident in the IB_Bias subsample (i.e. sample of firms in which IB issue positive recommendations which are not accompanied by a positive recommendation by the IND following the same company) than the IND_Bias subsample (i.e. the sub-sample of firms for which IND-analysts issue positive recommendations which are not confirmed by the IB-analysts following the same company).

The premise is that given greater incentive to ingratiate, IB-analysts' reports in these circumstances will be more optimistic in tone and less readable than IND reports. Table 40 and 41 present the results for the paired sample and related conditional logistic regression.

Table 40: Paired sample test of difference in linguistic measures – IB positively biased recommendations

Panel A: Difference in mean values

Variable	IB	IND	Diff
Bog	69.69	56.06	13.63***
Tone	2.87	2.10	0.78**

Panel B: Difference in median values

Variable	IB	IND	Diff
Bog	71.5	51	20.5***
Tone	3.03	2.00	1.03**

This table shows the results of a paired sample test of difference in mean and median scores of the measures of tone and readability between IB and IND-reports for a sub-sample of firms for which IB-analysts' recommendations are relatively optimistic (i.e. positive IB-recommendation accompanied by a negative or neutral IND-recommendation for the same companies). Panel A reports test of difference in the mean values using the paired sample t-test, while Panel B reports test of difference in median values using the non-parametric equivalent of the t-test, the Wilcoxon signed rank test. The total number of observations is 64 (32 IB analysts' reports and 32 IND analysts' reports). *Bog* is the *Bog index* which is used to measure the readability of the reports in this study. Higher values indicate more syntactic complexity and thus less readable reports. *Tone* measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). ***/**/* means significance at 1%, 5% and 10% respectively.

Table 41: Paired sample difference in linguistic features - conditional logit - IB positively biased recommendations

IB	Coefficient	Odds Ratio	Z	P>z
Bog	0.09	1.10	2.71	0.01
Tone	0.94	2.55	2.09	0.04
N	64.			
Wald X^2	7.50			
Significance	0.02			
Pseudo R^2	0.46			

This table shows the results of using the conditional logistic regression model to test for differences in the linguistic features of IB and IND-analysts' reports for the IB_Bias sub-sample. The dependent variable is the logit of the odds of a report being issued by an IB-analyst. The variable *IB* takes the value of 1 for reports issued by an IB analyst and 0 for reports issued by an IND analyst. *Bog* is the *Bog index* which is used to measure the readability of the reports in this study. Higher values indicate more syntactic complexity and thus less readable reports. *Tone* measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). The odds ratio is the exponentiated coefficient. Z is the z-value based on the Wald test of significance of each of the coefficients. The significance of the overall model is tested by the Wald X^2, which test whether the coefficients are jointly equal to zero. ***/**/* means significance at 1%, 5% and 10% respectively.

The results in Tables 40 and 41 show that relative optimism in recommendations of IB-analysts is accompanied by relative optimism in the tone of the reports as well less readability compared with IND-analysts. The difference in readability as measured by the *Bog index* is positive and statistically significant at the 1% level.

Similarly, the difference in tone is positive and statistically significant at the 5% level. The conditional logistic model is also statistically significant at the 5% level with a Pseudo R^2 of 0.46. This suggests that a greater proportion of the variation in the reports of IB-analysts is explainable by the tone and readability of the reports for this sub-sample. The coefficients are positive and odd ratios greater than 1 and significant at the 5% level. This implies that a unit increase in the tone and readability variables increases the odds of the report by an IB-report by a factor of 2.55 and 1.10 respectively. In other words, IB-analysts' reports are more likely to be more complex and more optimistic which suggest a strategic reporting, particularly as the positive recommendations which accompany such reports are not supported by IND-analysts and is consistent with previous impression management studies and supports $H_{3A \text{ and } 3C}$.

The above tests are repeated for companies for which IND-analysts have issued a positive recommendation which is accompanied by a neutral or negative recommendation by IB-analysts. Given the premise in this study that IB-analysts are more incentivised to curry favour with corporate managers through optimistic research, Hypotheses H_{3B} and H_{3D} postulate that the tone and readability of IND analysts' reports will not be significantly higher than those of their IB colleagues. Results of paired sample tests and conditional logistic regression are reported in Tables 42 and 43.

Table 38: Paired sample test of difference in linguistic measures – IND positively biased recommendations

Panel A: Difference in mean values			
Variable	**IB**	**IND**	**Diff**
Bog	67.64	56.82	10.82***
Tone	2.01	2.75	-0.75**
Panel B: Difference in median values			
Variable	**IB**	**IND**	**Diff**
Bog	66	58	8.0***
Tone	1.87	2.73	-0.87**

This table shows the results of a paired sample test of difference in mean and median scores of the measures of tone and readability between IB and IND-reports for a sub-sample of companies for which IND-analysts' recommendations are relatively optimistic (i.e. positive IND recommendation accompanied by a negative or neutral IB-recommendation for the same companies). Panel A reports test of difference in the mean values using the paired sample t-test, while Panel B reports test of difference in median values using the non-parametric equivalent of the t-test, the Wilcoxon signed rank test. The total number of observations is 78 (39 IB analysts'

reports and 39 IND analysts' reports). *Bog* is the *Bog index* which is used to measure the readability of the reports in this study. Higher values indicate more syntactic complexity and thus less readable reports. *Tone* measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). ***/**/* means significance at 1%, 5% and 10% respectively.

Table 39: Paired sample difference in linguistic features - conditional logit - IND positively biased recommendations

IB	Coefficient	Odds Ratio	Z	P-value
BOG	0.06	1.06	2.47	0.01
Tone	-0.28	0.75	-1.60	0.11
N	78.00			
X^2	9.13			
Prob > X^2	0.01			
Pseudo R^2	0.26			

This table shows the results of using the conditional logistic regression model to test for differences in the linguistic features of IB and IND-analysts' reports for the IND_Bias sub-sample. The dependent variable is the logit of the odds of a report being issued by an IB-analyst. The variable *IB* takes the value of 1 for reports issued by an IB-analyst and 0 for reports issued by an IND-analyst. *Bog* is the *Bog index* which is used to measure the readability of the reports in this study. Higher values indicate more syntactic complexity and thus less readable reports. *Tone* measures the tone of the reports based on the combined dictionaries of Loughran and McDonald (2011) and Henry (2008). The odds ratio is the exponentiated coefficient. Z is the z-value based on the Wald test of significance of each of the coefficients. The significance of the overall model is tested by the Wald X^2, which test whether the coefficients are jointly equal to zero. ***/**/* means significance at 1%, 5% and 10% respectively.

The results in Tables 42 and 43 show that for the sub-sample of firms for which IND-analysts issue relatively optimistic recommendations compared to IB-analysts, the accompanying narratives are not less readable. The difference in readability between IB and IND-analysts is still positive and significant at the 1% level (both for the paired sample t-test and Wilcoxon signed test of differences in median). The result is also maintained using the conditional logistic regression as reported in Table 43. The coefficient on the *Bog index* variables is positive and significant at the 5% level. The odds of issuing a less readable report are still higher for IB-analysts.

The difference in the tone of the reports for this sub-sample is negative with a mean score for the tone variable of 2.75 for IND-analysts compared with a 2.01 for IB-analysts. The paired sample t-test and the Wilcoxon sign test show that the difference in tone is significant at the 5% level. However, the result is not maintained under the conditional logistic model as the coefficient on the tone variable is not statistically significant. Overall the results lend support to

hypotheses H_{3B} and H_{3D} and reveal that strategic reporting in the form of optimistic tone and syntactic complexity doesn't characterise IND-analysts' reports when their recommendation is relatively more positive. However, this is not the case for the IB-analysts, whose relatively positive recommendations are accompanied by strategic reporting in the form of higher optimistic tone and more syntactic complexity. Malmendier and Shantikumar (2014) suggest that where optimistic bias is motivated by incentives, analysts use both recommendations and earnings' forecasts to achieve their goals. Similarly, the present study reveal that analysts' more incentivised to curry favour with corporate managers through optimistic research do not stop at issuing more positive recommendations but also use the accompanying narratives strategically.

10.4.4 Hypotheses $H_{4A \text{ and } 4B}$: Association between optimism and readability

H_{4A} and H_{4B} test the extent to which the level of optimism reported by analysts is associated with more complex reporting. It is anticipated that strategic reporting will be evidenced in reports in which pessimism is associated with less readable reports and optimism is associated with more readable reports. Hence, the overall sample is split into IB sub-sample and IND sub-sample. The aim here is not to test the extent of difference in the association between both sub-samples, rather the interest is to estimate whether strategic reporting pattern is present in the IB sub-sample as opposed to the IND sub-sample. The results of the multivariate analysis are presented in Table 44. The main variable of interest is the tone of the reports and other variables are included for control purposes.

Table 40: OLS regression results of optimism and readability

	Bog_IB			Bog_IND	
	Coeff.	**T-values**		**Coeff.**	**T-values**
Constant	89.41	5.73***		71.43	5.85***
Tone	-2.12	-2.77**		0.14	0.21
P_EPS	0.30	0.64		-0.17	-0.44
Size	-1.07	-0.69		-1.00	-0.61
Leverage	-20.88	-4.14***		-5.10	-0.50
Risk	-53.33	-1.47		9.12	0.39
Growth	5.46	0.53		40.75	2.83**
Intan	10.65	1.36		0.17	0.03
MTB	0.07	4.12***		-0.05	-3.96***
LTC	-2.03	-0.78		-9.41	-4.78***
POS	-0.28	-0.11		-0.36	-0.10
N		132			128
F		1.83			4.13
P-value		0.06			0.00
R^2		0.10			0.23

This table shows the results of multivariate analysis of the *Bog index* regressed on tone and company characteristics which is aimed at testing the association between financial performance (measured by P_EPS) and tone. Two models were estimated. The first model is for the IB-analyst sub-sample and the second model is for the IND-analysts sub-sample. The dependent variables in the models is *Bog* which measures the readability of the reports using the *Bog index*. The dependent variables are: tone, P_EPS is the percentage change in EPS from prior year. Size is the natural logarithm of market value for the fiscal year end. Leverage is the long term debt to total asset ratio. Risk is the standard deviation of EBIT over five years scaled by average assets over the

time period. Growth is the cumulative annual growth rate in sales per share over the previous five years. Intangibles is the proportion of intangible assets over total assets. MTB is the ratio of market value of equity to book value of equity for the fiscal year end. POS is dummy variable which takes the value of 1 for reports with a positive recommendation and 0 otherwise. Firms are classified into two groups, i.e.hi-tech (HTC) and low-tech (LTC). Classification of firms into HTC and LTC is based on the classification scheme used in Francis and Schipper (1999). LTC is dummy variable which takes the value of 1 for reports for firms in the low-tech industry and 0 otherwise. ***/**/* means significance at 1%, 5% and 10% respectively.

Both models are significant at the 1% level (IND) and 10% (IB) with an R^2 of 0.1 and 0.23 respectively. The variable of interest is *Tone*, while the others are control variables. The coefficient on *Tone* is negative and significant at the 5% level for the IB sub-sample and not significant for the IND sub-sample. First, this suggest that for the IB sub-sample, reports with higher tone (i.e. more optimistic) have lower *Bog index* values indicating more readability. Hence readability is positively associated with tone and lends support for hypothesis H_{4A}. This support the argument in Malmendier and Shantikumar (2014) that analysts incentivised to issue optimistic research, manipulate more than one of their outputs. In this case, IB-analysts use both the readability and level of optimism jointly as they issue more readable reports when they are more optimistic and less readable reports when they are pessimistic about the company. The coefficient for the tone variable is not significant for the IND-model and reveal that IND analysts do not strategically vary the readability of their reports as the level of optimism of their reports vary, providing support for H_{4B}.

10.4.5 Discussion of results

Overall, the findings suggest that the narrative content of analysts' reports are important when examining analysts' bias. The main difference between IB and IND-linguistic features are in the use of complex reports (i.e. reports with low readability). As previously documented in prior studies (such as Li, 2008), increased complexity is strategically used to obfuscate bad news. IB-analysts' strategy of reporting bias consists principally in the use of more complex reports relative to IND-analysts and the results in Table 44 reveal that the variation in the readability of IB-analysts' reports is systematically associated with the level of optimism reported as measured by the tone. Their reports are more readable when optimistic and less readable when pessimistic. The pattern is not observed for IND-analysts which are arguably (or at least perceived to be) more objective.

The general pattern is that the tone of IB-reports is not significantly different from those of IND-reports following the same firms. However, when the sample is limited to firms for which IB-analysts have issued relatively positive recommendations, both the tone and readability of the reports are significantly higher than the IND-reports. The same pattern is not observed for IND-analysts. Moreover, the tone of IB-analysts is less associated with changes in widely cited measures of financial performance such as EPS compared with the tone of IND-reports. It could be argued that IB-analysts, having greater resources, have more access to other qualitative information beyond financial statements data such as EPS (Clarke *et al.*, 2011) and could rely more on this information when writing their reports. However, Penman (2003) observes this pattern during the dotcom bubble but suggests that ignoring fundamentals such as EPS relative to innovative measures of performance actually fuelled the bubble leading to the subsequent financial crisis. Previous studies also show that analysts include and discuss EPS measures much more than any other financial information (e.g. Previts *et al.*, 1994). This was further supported from the manual content analysis conducted for Part A of this thesis. Thus, it is expected that the overall tone of the reports should be significantly associated with such widely used measure of company performance. Taking together, these results suggest that IB-analysts may engage in subtle manipulation of the syntactic content of their reports consistent with their motivations to issue optimistic research as a means of ingratiating corporate managers. Given the premise that bias motivated by incentives is perceived as an

ingratiatory behaviour aimed at eliciting rewards from managers, this study shows that the reporting pattern of IB-analysts who are more incentivised to curry favour (i.e. ingratiate) with corporate managers differ from those of IND-analysts in ways that suggest strategic reporting.

The findings extend previous studies on analysts' bias which have largely ignored the narratives which accompany analysts' reports. The evidence of some pattern of strategic reporting by IB-analysts reveals that the narratives which accompany the recommendations are useful in examining analysts' bias. It is also particularly relevant for policy makers who have chiefly focused on analysts' recommendations in the course of legislating on the issue of analysts' bias. More academic research is necessary to shed light on how narratives are used to present an optimistic view of firms and their managers by analysts. Interesting avenue for future research include the study of selectivity bias, other forms of concealment such as omission of relevant but negative news, analysis of attributional statements, the use of visuals, forward looking information, self-references, passivity of sentences, among others.

10.5 Sensitivity analysis

To assess the validity of the results, several checks were carried out and the sensitivity of the results to alternative measures or tests were conducted. Some of the results have been discussed in earlier sections. In this section, these checks are reviewed and summarised.

First, alternative measures of tone and readability were used in the paired sample tests. For tone, the results for the paired sample t-test were further analysed using a measure of tone based only the L&M wordlist as this has been used by previous studies of analysts' reports (Twedt and Rees, 2012) and for readability, popular measures such as the Fog and Flesch-Kincaid measures were used. Untabulated results show the results are robust to alternative measure of tone. However, for the readability measures, results are mostly confirmed using the Flesch-Kincaid measures while the Fog index yielded opposite results. Through the three formulae are based on the sentence length and word difficulty, the *Bog index* rates difficult words differently from the other two and also includes additional penalties for acronyms and other style problems which are highlighted in the SEC's plain English requirement for corporate documents as previously discussed in Chapter 9 of this thesis. Thus, it provides a conceptually superior measure of readability than the other two measures. Moreover, the different results between the Fog and the Flesch-Kincaid measures reveal that these variables measure readability with different level of noise.

Second, as previously mentioned in previous chapters, the sample of reports used in this study is matched between IB and IND-analysts. Cram *et al*. (2009) suggest that, for choice-based and/or matched samples, the results may not be generalised for two reasons: first, choice-based samples are non-random and, second, the number of observations in the sample groups and corresponding matches are not reflective of the sizes of the corresponding groups in the population. As previously highlighted in Chapter 6, disproportionate sampling occurs in this study as the incidence of IND-reports is significantly lower than those of IB-reports in the population of reports available on *Investext*. Hence, the robustness of the paired sample t-test reported was further assessed using conditional logistic regression and results are presented and discussed in previous

sections. As highlighted in that section, the results of the paired sample t-test are maintained with the use of conditional logistic regression.

Third, the standard errors for the regression models reported in this chapter are based on the White (1980) heteroscedastic-robust standard errors and thus controls for possible heteroskedastic concern in the data.
Fourth, the robustness of the analysis to clustering (as previously tested for the results in Chapter 6) was also conducted. Untabulated results reveal that the findings are largely consistent across models after controlling for differences in writing styles across analyst-firms.

Finally, multicollinearity checks were conducted for each of the regression models by computing the VIF scores. According to Wooldridge (2013), VIF scores greater than 10 indicate likely presence of multicollinearity. For each of the models, all the VIF scores were below the value of 2. Hence, multicollinearity was not a concern for the analysis.

Chapter 11: Thesis summary, contribution, implication and future research

11.1 Introduction

This chapter summarises and concludes this thesis. The chapter begins by presenting an overview and summary of this research project in section 11.2, which includes research objectives, hypotheses and findings. This is followed in section 11.3 by a discussion of the contributions of the study. Next, the implications of the research for policy and practice are discussed in section 11.4 while the research limitations are highlighted in section 11.5. Finally, future research opportunities and extension of this research are discussed in section 11.6. Section 11.7 concludes the chapter.

11.2 Research overview and summary

The aim of this thesis has been to investigate the factors which explain the variation in the content of sell-side analysts' reports. In order to address this, two related yet distinct research objectives were pursued. The first research objective was concerned about the thematic content of the reports and was focused on investigating the factors which explain the variation in the use of accounting information in the reports. The second research objective was concerned with the linguistic features of the reports and was aimed at examining the factors which explain the variation in the tone and readability of the reports. The findings for each of the objectives are discussed separately in section 11.2.1 and section 11.2.2 respectively.

11.2.1 Overview and summary of the first research objective

The overarching research question for the first research objective was: What factors explain the variation in the use of accounting information in analysts' reports? To address this question, two sets of literature were reviewed: first, the literature which directly investigates use of accounting information by analysts; second, the value relevance literature. The former provides a context for this research as it directly investigates use of accounting information by analysts while the latter literature was relied on to provide theoretical justification for the role

of accounting information in equity valuation, which is a key task of analysts' intermediation process. To address the research question, hypotheses were formulated which test the association between the extent of use of accounting information by analysts and a variety of company characteristics, previously studied in the value relevance literature. Further, hypotheses based on analysts' and report-characteristics were further specified. The measure of the extent of use of accounting information in analysts' reports was derived through content analysis of a sample of 288 analysts' reports for 144 S&P 500 companies and is the dependent variable. Using a detailed coding instrument, the reports were analysed and coded across two dimensions: the main topic, i.e. accounting information and the time-orientation of the accounting statements.

The empirical findings show that analysts use a wide variety of accounting information in their reports, although as previously documented in prior research, financial performance related information dominates. Lower level categorisation scheme used in this research provided additional information as it revealed that earnings is not the only or most important information referred to in analysts' reports as there were more references to revenue related information. However, this did not translate into greater use of revenue based valuation models, which suggested that relevant information for analysts differs from the traditionally construed notion in value relevance literature of relevant information being equated with price-sensitive information or information used as direct input into valuation models. Moreover, most references to accounting information are historical and reading of the reports reveal that historical references is mostly used to assess managerial performance with prior estimates and forecasts.

In general, the extent of use of accounting information varies significantly across reports. Examination of the sources of variation suggests that use of accounting information is partly explainable by company, analyst and report-characteristics. With regards to company characteristics, analysts' use of accounting information (both at the high and lower level categories) were found to be lesser, the greater disconnect between accounting information and company value (such as for firms with higher market-to-book ratios and firms in high-tech industries) and the lesser the level of leverage. Extent of use of financial performance related information (particularly historical references to profitability ratios) was positively associated

with changes in EPS, indicating that analysts use more financial performance information in their reports, the better the financial performance of the company. With regards to analysts' characteristics, analysts' use of accounting information (particularly forward-looking information) is higher for IB-analysts compared to IND-analysts and for CFA qualified analysts compared to non-qualified analysts. However, the latter results hold mostly for IND-analysts. The extent of use of accounting information is not associated with types of recommendation in most cases but is highly related to the length of the reports as measured by the word count.

11.2.2 Overview and summary of the second research objective

The second research question formulated to address the second research objective was: What factors explain the variation in the linguistic features of analysts' reports? This second objective goes beyond the accounting themes to examine variation in "how" information is communicated in analysts' reports. Three literatures were relied on to provide a context for this research objective. First, the literature on the linguistic features of analysts' reports was reviewed. Research in this area is at the early stages with only few studies, the majority of which examine the effect of linguistic features on market variables such as market returns. Despite evidence from research on corporate reports, which show that manager's tendencies to present a positive image of their companies is associated with strategic manipulation of the linguistic features of their reports, research is yet to examine whether analysts' tendencies to produce optimistic research about the companies they cover is associated with strategic manipulation of their reports. Hence, the second literature reviews the evidence with regards to analysts' optimistic bias. This literature has mostly focused on examination of bias by studying the summary output of analysts' research activities, i.e. earnings' forecasts, recommendations and price targets. The impression management theory of ingratiation was relied on to formulate testable hypotheses, which is the third literature reviewed. Hypotheses were mostly centred on the differences in linguistic features between IB and IND-analysts' reports.

Empirical evidence reveals that both types of analysts issue reports with optimistic tone and there is no significant difference in tone between IB and IND-analysts while the readability of the reports issued by IB-analysts is lower compared with

IND-analysts. However, when the sample is restricted to cases of relative optimism in recommendations by IB-analysts, the tone of their report is higher while the readability is lower compared with IND-analysts. Similar patterns were not observed when the sample is restricted to cases in which IND-analysts are relatively more optimistic in their recommendations. This suggests that IB-analysts' optimism in recommendation is accompanied by strategic reporting. Moreover, when the sample is restricted to cases of negative changes in financial performance (as measured by EPS), there is no significant difference in tone but the readability of IB-reports is lower than IND-reports. Further test of association between tone and financial performance revealed that the tone of IND-reports is significantly and positively associated with changes in financial performance. While, a similar result is observed for IB-reports, the magnitude is lower. Finally, the results reveal that for the IB-analysts, the measure of readability of their reports is negatively associated with the tone of the reports, which further suggests a strategic variation of readability as the tone of the reports varies, indicative of obfuscation tendencies

11.3 Research contribution

The research makes a number of contributions to the academic literature empirically, theoretically and methodologically.

Empirical Contribution

The empirical findings from this thesis provide an extension to several branches of literature. First, this study contributes to the literature on the content of analysts' reports in several ways: It provides evidence on the disclosure of accounting information in analysts' reports, the extent of variation of accounting information across analysts' reports and the factors which explain such variation. Previous studies of accounting information have used broad categorisation schemes which include accounting information in either a financial/operating category (Rogers and Grant, 1997; Nielsen 2007), or have only captured a subset of accounting information such as earnings' quality (Barker and Imam, 2008). The present study provides clear and comprehensive evidence on the use of accounting information both at the high level classification (of accounting information into financial performance and financial position related information), at lower levels

(e.g. distinguishing financial performance into earnings, revenue and expense related information) and the time-orientation of the references to accounting information, distinguishing between historical, forward-looking and non-time specific references. More specifically, this study extends existing papers on usage of accounting information by analysts as it provides evidence on how the use of accounting information varies across analysts' reports and identified the company, analysts and report-characteristics which influence such variation. Evidence of analysts' use of accounting information and the factors which influence their usage is important as it provides input into the debate about the relevance of accounting information for users, which is of particular interest to standard setters, accounting professional bodies, corporate managers and accounting academics. Overall, the evidence suggests that usage of accounting information is not driven primarily by valuation concerns or company characteristics. User's characteristics such as their incentives, also play an important role in shaping their choice of accounting information.

Second, this study extends the literature on the value relevance of accounting information. The value relevance literature has traditionally presented earnings as the most relevant financial performance measure for equity valuation (Barton *et al*. 2010). However, this study documents the usefulness of other types of accounting metrics such as revenue, which was found to be most cited financial performance metric in analysts' reports. Thus, the findings here support the inclusion of revenue in future value relevance studies as was examined in Chandra and Ro (2008).

Third, this study extends the literature on the linguistic features of analysts' reports. In general, previous studies have either focused on the description of these features or examined their associations with market data. This study extends the literature by investigating variations in the linguistic features and examining whether these are associated with analysts' incentives to issue optimistic research. Hence, , this study contributes to the literature on analysts' bias as it provides evidence that optimism is not limited to summary measures such as recommendations, forecasts or price target but may extend to the narrative content of analysts' reports. The evidence that analysts' reports are generally optimistic, containing more positive than negative expressions support

the existing research which document the tendency of analysts to issue optimistic earnings' forecasts and more positive recommendations. Further, the empirical evidence suggests that differences between IB and IND-analysts are not limited to the summary measures which accompany their reports as there are also different pattern of reporting. This evidence is important as it improves our understanding of how analysts' incentives to issue optimistic research may affect other research output such as their written analysis. This is important both for investors who rely on these reports as well as policy makers seeking to address analysts' bias.

Theoretical Contribution

From a theoretical perspective, this study extends the literature on analysts in general by introducing the impression management theory of ingratiation. The term "curry favour" is synonymous with ingratiation and is widely used to describe analysts' relationship with corporate managers in existing research on analysts' bias. Nevertheless, a well-structured theory has been elusive. Some studies (e.g. Barker and Imam, 2008; Abhayawansa and Guthrie 2012) have also referred generically to impression management to explain findings which relates to the content of analysts' reports. This thesis builds on the study of Westphal and Clement (2008) and introduces the impression management theory of ingratiation which explains the social relationship which exists between corporate managers and analysts and provides a relevant theoretical model for explaining analysts' impression management tendencies.

Methodological Contribution

This study provides a number of methodological contributions to existing literature on the use of accounting information in analysts' reports. First, the present study provides a clear and comprehensive working definition of accounting information. The classification scheme includes lower level categorisation of accounting information and considers the time-orientation of accounting statements in the report. Such details enabled analysis not featured in extant literature and provide additional insight into the types of accounting information used by analysts and how it is used. The definition of accounting information, categories of accounting information and coding rules developed for this study could be used in future

studies of accounting information in analysts' reports or other financial documents.

A second methodological contribution to the literature on the content of analysts' reports is the use of a recording and measuring unit which captures the space that accounting information occupies in the reports, while adhering to principle of mutual exclusivity and enabling a study of variation of the content of the reports. Hence, future studies aimed at investigating variation in the reports could adopt the recording and measuring units used to address the first research objective.

A third methodological contribution is the use of a combined wordlist developed by Henry (2008) and Loughran and McDonald (2011) to measure the tone of analysts' reports. Previous studies have largely relied on wordlists in text analysis software such as *Diction* or *General Inquiry* (Fogarty and Rogers, 2005; Mokoaleli-Mokotelli *et al*. 2009; Twedt and Rees, 2012). However, as demonstrated in this present study, these wordlists are not very suitable for analysis of the tone of analysts' reports. For instance, misclassification errors based on the use of the wordlist in *Diction* are large, which questions the validity of previous studies based on the use of that software. Following checks of type I and type II errors, the wordlists used in this study were found to produce a more valid measure of tone in analysts' reports. This combined list could be used in future research of the tone of analysts' reports in place of standard wordlists contained in *Diction* or other textual analysis software, which were developed from other subject areas such as politics and psychology.

Fourth, unlike previous studies, a check of the validity of the findings using wordlists was conducted in this study. This involved checking the context of each tonal word to ascertain whether it was correctly classified into a positive or negative category (using a small sample of reports). The extent of misclassification was reported. Future studies of the tone of analysts' reports based on wordlists could conduct similar tests on a sub-sample of their data to provide readers with some confidence in the reported results.

Fifth, this study uses the *Bog index* to measure the readability of analysts' reports, unlike previous studies which used traditional measures of readability such as the *Fog index*. As explained in Chapter 9, the *Bog index* offers a valid alternative to

the traditional measures of readability such as the *Fog index*. Future studies of analysts' reports could use the *Bog index* to measure readability as was used in this thesis and is currently being used to study corporate reports (E.g. Bonsall *et al*. 2015).

11.4 Implications of research findings

The findings reported in this study have relevant policy implications. First the evidence suggest that earnings and profit measures are not the most important financial information metrics used by analysts. Though references to revenues in analysts' reports are pervasive, they are not used as direct input into traditional equity valuation models. Hence, the implication is that debate about relevance of accounting information by standard setters or the accounting profession should go beyond its conceptualisation in the value relevance literature which is traditionally focused on earnings and based on price-sensitivity (Barker *et al*. 2012; Abhayawansa *et al*. 2015). Considerations should be given to uses of accounting information from users' perspectives such as its use for evaluation of other types of information. For instance, Bricker *et al*. (1995) and Barker and Imam (2008) reveal that the quality of accounting earnings (which is directly used in equity valuation models) is based on the assessment of other accounting information such as revenue. Moreover, the influence of user's characteristics on the relevance of accounting information should be considered. This study revealed that user's incentives and training play an important role.

Equity analysts are largely concerned with gathering information to estimate the value of the companies. Hence, the finding that most references to accounting information are historical (consisting mostly of comparison of company performance with previous estimates and benchmarks) rather than forward-looking also imply that accounting standard setters should re-evaluate whether stewardship and valuation role are alternative roles of accounting information or two sub-components of the broader concept of decision usefulness (Gassen, 2008)

Additionally, the findings from the second research objective which showed differences in readability of IND and IB-analysts' reports provide input into policy intervention with regards to analysts' bias. First, it provides justification for the requirement of IND-analysts' reports as part of the Global Settlement Act as this

research shows that IND-analysts' reports are more readable compared to those of IB-analysts' reports. Second, evidence of strategic reporting also provides relevant information for future policy intervention on analysts' bias, which should extend beyond the summary measures such as recommendations.

The findings reported in this research are also relevant for corporate managers, who produce accounting and other information for intermediaries and investors. First, the evidence that analysts' use of accounting information vary across company characteristics imply that corporate managers should consider the features of their companies when deciding on the provision of voluntary information. For instance, this study reveals that historical accounting information is less used by analysts following companies in high-tech industries. Hence, managers of companies in such industries should provide more forward-looking information. Second, the variation in use of accounting information, tone and readability of analysts' reports imply that corporate managers should be aware of the several ways in which analysts' discretionary choices (of information or the form of presentation) may influence the perception of their companies in capital markets. Beunza and Garud (2007) explains analysts' role in this regard as similar to frame-makers who provide a frame through which investors value stocks during uncertainty.

For investors who rely on the content of analysts' reports, the finding that analysts' use of accounting information varies significantly across analysts' reports and is influenced by analysts' characteristics (such as whether they are employed by IB-firms or IND-firms and whether they are CFA qualified) suggest that investors should rely on more than one analysts' reports in forming their opinions about the companies they invest in. Additionally, in selecting reports, these analysts' features should be considered. For instance, reports by IB and IND-analysts or reports by CFA and non-CFA chartered analysts should be examined as these attributes are relevant in explaining variation in the content of analysts' reports. Moreover, investors should be conscious of analysts' discretion in the inclusion of information in their reports as different analysts' following the same company may choose and comment on information differently. Moreover, as the findings of this study only partly support the valuation rationale for the use of information, investors should be conscious of other possible reasons why use of information

may vary such as the need to sell a story about the companies they report on. Finally, the need to consult a wide range of analysts' reports and be conscious of analysts' incentives is also necessitated by the evidence from the second part of the research, which reveals that IB-analysts' reports are more complex and associated with the overall level of tone of their reports, compared with IND-analysts.

Finally, there are a number of implications for future research studies. First, the lower-level categorisation and consideration of the time-orientation in this study provide additional insight into use of accounting information. Hence, future studies should consider other dimensions of reporting as previously advocated in Beattie *et al.* (2004). Second, the evidence that analysts' use of accounting information varies significantly across IB and IND-analysts imply that academic research on the difference in recommendation and earnings' forecasts between these analysts have only provided a one-sided evidence as the differences in the output may actually be influenced by differences in the input (information) used. Hence, future studies aimed at investigating the difference between IB and IND-analysts should go beyond their output. Consistent with this argument, the second research objective also revealed that the linguistic features of analysts' reports partly differ across these analysts-characteristics. Moreover, the evidence from this research is useful in shaping future studies of analysts' reports' linguistic features. For instance, in examining the information content of the linguistic features, researchers might need to control for analysts' incentives as market reaction may differ between IB and IND-analysts' reports.

11.5 Limitations

While the research has been designed to enhance reliability and validity of the overall findings, it is still subject to potential limitations as follows:

First, the content analysis approach used in this study is based on quantification and frequency of mention of accounting information items is taken to be indicative of its relevance. However, analysts' reports are not record of analysts' decision making processes (Govindarajan, 1980). Hence, choice of information used may be different from information disclosed in the reports.

Second, manual content analysis is inherently subjective. While detailed and well-specified coding schemes enhance reliability, it does not completely address it. Steps were taken to enhance objectivity and reliability of the results as discussed in chapter 5 of this thesis and inter-coder and intra-coder reliability tests are generally of an acceptable level.

Third, the use of wordlists to measure the tone of analysts' reports enhances reliability but may not produce measures which are completely valid. To address this limitation, a pilot sample was used to test the validity of alternative wordlists. The research is based on the use of a wordlists which provide lower misclassification errors. Nevertheless, the use of wordlists ignores the context in which words are used and some words derive their tonal meaning from the context.

Fourth, this study is based on the analysis of results reports only. There are several types of reports issued by analysts as previously discussed in this thesis. However, the choice to focus on this type of reports is driven principally by lack of availability of other types of reports for both IB and IND-analysts.

11.6 Future research

There are several avenues for future research.

First, this study revealed that accounting information about financial position are rarely discussed in analysts' reports. One of the rationales for this is that equity research reports are generally focused on a company's equity value, with companies generally evaluated on a going-concern basis. Hence, to provide a more comprehensive empirical evidence of the use of accounting information for company valuation, future research may complement the findings reported here by examining the use of accounting information in debt-analysts' reports. Beyer et al. (2010, p.335) argued that "research on the interplay between information provided by sell-side security analysts, other information intermediaries such as debt analysts, and firms' mandatory and voluntary disclosures is warranted". Hence, differences in the use of accounting information by debt and equity analysts may further complement the findings from this research in response to this call for further research.

Second, future studies may extend the findings in this research by using qualitative content analysis methods which are not based on quantification. This is likely to enhance understanding of how analysts use information.

Third, this research could be complemented by interview based study aimed at understanding how analysts make decisions on which information to disclose in their reports and whether there is a discrepancy between information used for valuation and that disclosed in their reports.

Fourth, future research could also investigate the interaction between impression management and value relevance of analysts' output.

Fifth, future research could extend this study to examination of other types of analysts' reports such as initiating coverage reports, reports issued prior or after relevant company events such as initial public offerings, seasoned equity offerings, mergers and acquisitions etc.

Sixth, this research was based on a US sample. Future research may extend this study to other countries to examine whether reporting practices differ across analysts' in different jurisdictions.

Seventh, future studies of the tone of analysts' reports may extend this study by developing a word lists from a large sample of analysts' reports which could be used in studies of the tone of analysts' reports. The lists used in this study were derived from earnings' press releases (Henry, 2008) and annual reports (Loughran and McDonald, 2011). There is currently no wordlists developed from analysts' reports, which implies that tonal words commonly used in these documents may not be fully reflected in existing wordlists. Hence, there is a need for future research to develop wordlists from a large sample of analysts' reports which could be used in future studies of their syntactic content.

Eighth, this study has extended the study of analysts' bias to the narratives which accompany their reports. However, only the variation in the tone and readability of their reports was examined. Future studies might extend the findings in this study by examining other impression management tactics such as attribution

patterns, selectivity bias and other strategic reporting practices studied in corporate reporting setting.

Ninth, ingratiatory theory suggests that both situational and dispositional factors influence ingratiatory behaviour. This study has focused only on one situational factor, i.e. analysts' incentives proxy by whether they are employed by an IND or IB-firms. Future research may also examine differences in readability and tone based on other factors. Other situational factors may include corporate events such as Mergers and Acquisitions or distinction between analysts employed by brokerage firms with existing relationship with companies (affiliated) and those without (unaffiliated). Dispositional factors may include analysts' own reputation as measured by their rankings in institutional investor surveys.

Appendices

Appendix A – Sample composition of analysts' firms

This appendix provides details of the analysts-firms, whose reports are used in this study. For each firm, details of the number of reports within the sample and the percentage of the overall sample are reported.

Analysts' firms	Number of reports	Percentage of overall sample
ACI Research	3	1.04%
Argus	13	4.51%
Auriga	1	0.35%
Barclays Capital	2	0.69%
Barrington Research	3	1.04%
BB&T Capital Markets	1	0.35%
Blue Capital	1	0.35%
BMO Capital Markets	3	1.04%
Buckingham Research Group	8	2.78%
Canaccord Genuity	4	1.39%
Capstone Investments	1	0.35%
Caris & Company	2	0.69%
CL King & Associates	1	0.35%
Collins Stewart	3	1.04%
Credit Suisse	9	3.13%
Deutsche Bank	8	2.78%
FBN Securities	1	0.35%
Guggenheim	1	0.35%
Hilliard Lyons	2	0.69%
HSBC Global Research	1	0.35%
Indigo Equity Research	12	4.17%
J.P. Morgan	6	2.08%
Janney Capital Markets	4	1.39%
Jeffries	5	1.74%
Kaufman Bros	1	0.35%
Keybanc Capital Markets	4	1.39%
Macquarie	7	2.43%
Morgan Stanley	15	5.21%
Morningstar	58	20.14%
Murphy & Durieu LP	2	0.69%
Northcoast Research	12	4.17%
Oppenheimer	7	2.43%
Piper Jaffray	4	1.39%
Pivotal Research Group	4	1.39%
RBC Capital Markets	9	3.13%
Roth Capital Partners	1	0.35%
Singular Equity Research	1	0.35%
Societe' Generale	1	0.35%
Stephens	1	0.35%
Suntrust Robinson Humphrey	1	0.35%
Susquehanna Financial Group	2	0.69%
Think Equity LLP	3	1.04%
Wall Street	40	13.89%
Wedbush	4	1.39%
Wells Fargo Securities	11	3.82%
William Blair	4	1.39%
Williams Capital Group	1	0.35%
Total	**288**	**100%**

Appendix B: Coding scheme

This appendix contains the coding instrument used for section A of this thesis. It presents the main categories to which units of analysts reports were classified. Each of the high level categories are made up of several accounting key words which do not form a separate category on their own. For instance, the category 'Expense' contains keywords such as SG&A, operating expenses etc. The coding scheme below illustrates the coding rules for coding text into the listed categories. As suggested by Boyatzis (1998), for each category, there are: i) Label (code); ii) Definition; iii) Description of when the category occurs (flag); iv) Qualification or exclusion and v) Examples. Further information about the process of developing the coding scheme is discussed in the main Thesis.
Table 45 presents the coding instrument and is followed by Table 46 which contains additional instructions on coding. Table 47 list frequently used acronymns by analysts' in their reports. The meanings and corresponding category are also highlighted to aid the coding process.

Table 41: Coding instrument

Accounting Category	Code	Definition	Flags, Qualifications and Exclusions	Examples
Financial Performance	REV	Revenue	Reference to revenues (also abbreviated as 'revs' in some cases), turnover, sales, like-for-like, comparable same store sales (comps), top-line, revenue growth rate, For healthcare firms, reference to healthcare premium Reference to franchise and incentive fees received, e.g Marriot group **Exclusions:** Income from sale of assets or investments (included under 'Asset' category Reference to units sold Reference to selling price References to sales as an activity or selling at discounts References to revenue mix	For the fiscal year, sales increased 9.6% over the prior year. (Morningstar on AZO, 2011) For the full-year, sales grew 8% to $10,184.6 million (Northcoast on DF, 2011) Management expects FY12 revenue to grow in the range of 6% to 8%, in constant currency. (JPM on CA, 2011)

	EXP	Expense	Reference to expenses, expenditure or charges (if context suggests it relates to expense item), costs, COGS, SG&A, depreciation or amortisation, overheads, pension related expenses, spending (if accompanied by value of the expenditure). Include expense growth rate, cost structure, tax related expenses and rates, bad debts, interest expenses Include references to cost savings. **Exclusions:** Reference to dividend payments (included under 'Equity' category) References to capital expenditures - Capex (Included under 'Asset' category). Reference to non-corporate tax references e.g. tax on consumption. Use of word "expense" not in reference to costs or reference to "cost" not in relation to value of expenses: "….at the expense of" or "…cost effective," or "cost pass through"	Operating expenses decreased 1.9% to $650.1 million in the 4QF10. (Northcoast on DF, 2011) Depreciation and amortization expenses increased 5.2% to $387MM (vs. $368MM LY), while rent expenses grew 1.6% y/y to $149MM (vs. $146MM a year ago). (Deutsche Bank on KR, 2011) Input costs are a large component of ConAgra's cost structure. (Morningstar on CAG, 2011)

	EARN	Earnings	Reference to earnings, profits, losses (as negative earnings), EPS, EBIT, EBITDA Reference to bottom-line (when used in the context of earnings), income/net income/operating income, when used in context of earnings. Look out for acronyms such as OIBDA, OIBA Earnings stream **Exclusions:** Use of earnings as an adjective, e.g. in "earnings' call or earnings' announcements or release, earning power, earnings season	Additionally, we are introducing our FY2013 EPS estimate of $2.32. (Northcoast on CTAS, 2011). [….] which pushed Q4 GAAP results to a loss of $40 million (Blue Capital on VRSN, 2011). EBITDA appears to have bottomed. (Janney on DF, 2011) On the bottom line, D.R. Horton reported net income of $35.7 million. (WallSS on DHI, 2011)
	PRATIO	Profitability Ratios	Reference to profit related ratios such as margins (including operating margins, EBIT margins, profit margins, contribution margins). References to accounting returns such as ROE, ROI, ROIC. References to expense or operating ratios e.g. SGA-sales **Exclusions:** Reference to EPS as this is included under Earnings category. Reference to dividends ratios e.g. Divided per share. These are included under 'Equity' category. Reference to price margins	Operating margins were +1%, 360bps better than our -2.6% estimate. (Deutsche Bank on LEN, 2011) AutoZone generates great ROICs that have continually trended up over the last decade. (Morningstar on AZO, 2011) SG&A margins leveraged 77bp to 16.20% from 16.98% last year, significantly better than our estimate of a 35bp drop to 16.63%. (Deutsche Bank on KR, 2011)
	ACT	Activity Ratios	Reference to activity ratios such as : inventory turnover, inventory days, days sales outstanding (DSO), inventory/sales ratio etc	DSO decreased 1 day to 23 days. (ROTH on INTC, 2011) Inventory turnover reached a record level of 5.6-times. (Argus on JWN, 2011)

	OFPERF	Other Fin. Performance	Reference to non-specific income statement items such as one-off items, other income, interest income, profitability (or profitable), special items Reference to benefits or monetary gains (e.g. gains from sales of assets, derivatives) Reference to the income statement in general (except accompanied by a specific item in which case, code to the item category.) Other financial performance item not classifiable in any of the categories above, e.g. references to operating leverage **Exclusions:** Reference to share/price gains. Reference to market share gains.	Other income statement items did not vary substantially with our estimates. (ThinkEquity on AZO, 2011) Q4 reflects the benefit of a $27MM ($0.11) pre-tax gain (Jeffries on COL, 2011) The result mainly reflected an $0.08 benefit from other income (Wells Fargo on YUM, 2011)
Financial Position	ASS	Assets	Reference to the assets (cost/value) (tangible and intangibles e.g. goodwill), including asset write-down, impairments Reference to capital expenditure (Capex) or proceeds from asset sale Reference to cash balances and inventory Reference to any components of assets as classified in the balance sheet e.g. deferred revenue, receivables, short term investments etc. **Exclusion:** Simple description of the Assets such as legal ownership aspects or other qualitative attributes. References to cash movements/generation – these are coded as cash flows.	DVN likely closes on $3.2 billion sale of offshore Brazil assets to BP (Buckingham on DVN, 2011) Capex in the fourth quarter this year was $40.5 mln compared with $52.3 mln a year ago. (Susquehanna on ABC, 2011) The company ended 4Q10 with $1.7 billion in cash and cash equivalents and short-term investments. (Janney on GPS, 2011) The balance sheet lists the value of land, buildings and equipment at $2.3 billion. (Argus on JWN, 2011)

	LIAB	Liabilities	Reference to debts balances or changes thereof, bonds, loans, borrowings, liabilities, financial leverage or notes (when used in reference to debts) Reference to pension obligations, payables and other components of liabilities as classified in the balance sheet. **Exclusions:** Reference to 'liabilities' if context suggest not related to financial liabilities	We aren't concerned about ConAgra's debt balance. (Morningstar on CAG, 211) Accounts payables were up 8.7% to $4.23B from $3.89B twelve months ago. (Deutsche Bank on KR, 2011) Total debt outstanding was approximately $2.455 billion at the end of 2010, including $1.913 billion of long-term debt and $1.255 billion of debt due within five years. (Argus on TEG, 2011)
	EQ	Equity	Reference to book value of equity, capital (if used in the context of Equity) Reference to shares or stock (count or value). Reference to share buybacks or share repurchase programs (when accompanied by the value of the repurchase) Reference to Dividends/dividend pay-out ratio **Exclusions:** References to market value or share prices or share price gains or shares trading References to share repurchase programme which is not accompanied by the value of the repurchase	The firm has repurchased about $1.3 billion worth of shares during the past two years (Morningstar on NTAP, 2011) The board also approved a new 50 million share repurchase program (Argus on KMB, 2011) The share count decreased 1.1% to 375 million (Stephens on TSN, 2011) The company had $51.1 million in preferred stock outstanding at the end of 2010. (Argus on TEG, 2011) We forecast that Heinz will raise its shareholder dividend in the high single digits annually over the next five years (Morningstar on HNZ, 2011)

	LIQ	Liquidity Ratios	Reference to liquidity ratios such as current/quick ratio or liquidity in general	The company's current and quick ratio were about 1.7 (Morningstar on HRL, 2011) With a current ratio of approximately 1.9 and a quick ratio of 0.9 at the end of the third quarter (Morningstar on TSN,2011)
	SOL	Solvency Ratios	Reference to debt-equity, debt-asset, debt-capital, debt-earnings (e.g. EBITDA) or solvency in general Reference to equity ratio or leverage ratio References to interest coverage ratio	Total debt/EBITDA was less than 1.0 at the end of fiscal 2010 (Morningstar on BFB, 2011) At the end of 4Q10, PX's debt/capitalization ratio was 47.5%, compared to 47.2% at the end of 4Q09 (Argus on PX, 2011) We believe that an interest coverage ratio above 5.0 usually indicates healthy leverage (Argus on Yum, 2011)
	OFPOS	Other Financial Position	Reference to financial position items not clearly attributed to other categories e.g. reference to the 'balance sheet' or 'financial position' or 'capital structure' References to capital, working capital	Balance Sheet Continues to Improve (WellsFargo on RM, 2011) Patterson has a strong balance sheet and fairly safe capital structure. (Morningstar on PDCO, 2011) The decline was due to a smaller level of improvement in working capital (ARGUS on KMB, 2011)

Other Accounting Information	**VM**	**Valuation model**	Reference to accounting based valuation models such as: Earnings based: EPS multiple, EBITDA multiple, EBIT multiple, P/E (Price to earnings), P/EBITDA, PEG, earnings multiple Cash flow based: P/FCF, (Price to free cash flow), FCF yield, DCF (Discounted Cash Flow) model Book value based: P/B (Price-book), NAV (Net Asset Value) Dividend based: Divided yield, Divided discount model (DDM) Enterprise value based: EV/sales, EV/ EBITDA, EV/EBIT Sales based: P/S (Price to sales) Time orientation will depend on the period of the accounting info used. Multiples are often replaced by "x" or "times" **Exclusions** Only accounting valuation models are coded. So generic references to valuation multiples are ignored if not specifically related to accounting items.	Air Products trades at 13.8x 2012E EPS and 8.0x EBITDA (Jeffries on APD, 2011) We calculate our 13.2x group multiple using the dividend discount model Our DCF analysis supports a multiple in the 16-17x range (Deutsche Bank on JCI, 2011) On an EV/EBITDA basis, PepsiCo shares are trading below the peer group, at 9.2x our 2011 estimate. (HSBC on PEP, 2011) TEG shares are trading at 6.9-times our 2012 cash flow estimate of $7.15 per share (Argus on TEG, 2011) We assign a P/FCF multiple of 20.0x (prior 21.0x) to our 2011E FCF/share of $10.82 (Caris on AMZN, 2011) We value CVC using a DCF methodology (Pivotal on CVC, 2011) We believe that the current enterprise value of about 9.4-times trailing EBIT is fair (Argus on JWN, 2011)

	CF	Cash flows	Reference to cash flows such as free cash flow, operating cash flow, cash movements/generation/usage	Operating cash flow in fiscal 2011 was $37.5B vs. $18.6B in fiscal 2010. (JJB on AAPL, 2011) MKC generated cash from operations of $338 million in FY10 (Janney on MKC, 2011) The company generated US$430m in operating cash flow and US$269m in free cash flow (Macquarie on PLL)
	NSA	Non-Specific Accounting Information	Reference to accounting results, annual reports or accounting information not classifiable into other categories e.g. unspecified accruals Reference to accounting policy Reference to minority interest **Exclusions** When "results" are accompanies by specific accounting items, the statements should be coded to the accounting items rather than included in this category. References to results if context is not sufficient to determine whether accounting results are being referred to.	Kroger posted solid Q4 results (Deutsche Bank on KR, 2011) US financial results were in line with expectations (Wells Fargo on DTV) During 4Q10, Integrys changed its accounting policy for investment tax credits to better match the benefit of these credits with the cost of the related investments. (Argus on TEG)

Time Category	FL	Future Oriented	Use of present tenses or present continuous to refer to future expectations or events and prospects Use of future tense such as "will" Use of past tense but relates to a future event Discussion of expectations, possibility, potentials, intentions, opportunities or assumptions (using words such as "expect" "anticipate" "predict" "intend", "could" "should" "project" "possible" "likelihood") or visibility of accounting items Description of risk and potentials (including if-then scenario analysis) Forecasting of accounting items (using words such as "estimates" , "guidance" or "forward" "forecast" "potential" "target" "outlook") and use of past tense to refer to future estimates. Reference to specific future dates (from 2011/2012 onwards) or future time period (Next twelve months, long-term, over five years etc) For valuation models: use of forward looking accounting information (e.g. use of next years' earnings for multiples, forward or 'E' to indicate estimates) Description of future share repurchases with reference to current authorisation **Exclusions:** References to "estimates", "guidance" "target" if referred to prior accounting periods.	Here we project operating margins to trend toward the 16%-17% range over time (Morningstar on AMAT, 2011) Management provided FY11 capex guidance in a range of $1.7B - $1.9B (Deutsche Bank on KR, 2011) Further decreases in gross margin are possible (WallSS on KR, 2011) We do not expect the dividend to increase anytime soon (Argus on AEE) Revenue guidance for the fiscal year is expected to range between $4.0 and $4.1 billion (Northcoast on CTAS, 2011) Management initiated very weak F2012 EPS guidance of $5.75-5.85 (Cannacord on BDX, 2011) Alcoa is trading at 5.4x our 2012 EBITDA estimates (Macquarie on AA, 2011) With the shares trading at just 10.8x our new FY12 EPS estimate of $2.14 (Deutsche Bank on KR, 2011)

			Reference to “estimates”, “guidance” are not coded except accompanied with earnings or other accounting terms or can be discerned from context e.g. previous statement or subsequent statement.	
	HIST	Historical	Reference to a specific time period in the past (2010 or earlier) Use of past tense to indicate past events Use of present tense to indicate past information Discussion of current/ recent accounting results References to historical forecasts/estimates For valuation model: use of historic accounting information for valuation purposes (e.g. trailing EPS)	Despite $600+ million in cash outlays in Q4, Verisign closed out 2010 with $2.06 billion in cash (Blue Capital on VRSN, 2011) The company recently took on $1 billion of bank loans in India (Morningstar on QCom, 2011) The company still has a low debt/capital ratio (Morningstar on WAG, 2011) Q4 EPS of $1.45 exceeded our estimate (Piper on TGT, 2011) Gross margin continued to expand in the 4Q (Keybanc on TIF, 2011) Our target PE is calculated using the sector historical PEG (~1.4x) (Deustsche Bank on COV, 2011)

	NTS	Non Time Specific	Reference to on-going and continuous activities with both a past and future orientation (using words such as “continue”, “still”) Statements in which time period has not been clearly specified. Expressions of beliefs, opinions and statements of general facts or description of current situations	Balance Sheet Continues to Improve (Wells Fargo on RM, 2011) Linear continues to generate stellar profitability (Morningstar on LLTC, 2011) Amgen does not pay a dividend (Argus on AMGN, 2011) Microchip has one of the highest dividend yields in the chip sector. (Morningstar on MCHP, 2011). However, we do not believe debt/capital is a good indicator of financial health for this firm, Morningstar on AZO, 2011). We believe investors have largely banked on relatively stable gross margins (RBC on LOW, 2011)

Table 42: Additional coding instruction

Cases	Coding rule	Example
Reference to activity, model, program, management using accounting terms. E.g. "inventory management" "buyback program" etc	Where a value for the accounting term is included in the statement, this is to be coded as accounting.	Management announced a **$400-$700** million share buyback **program** in 2011.
Caveats (or words in parenthesis)	Where the caveat includes information about a different accounting keyword, then should be coded to a separate category	"EPS of $3.50-$3.60 (**excluding $0.20-$0.22 of FAS-123 expenses**)" (CERN)
Use of pronouns to discuss accounting keywords	If the keywords are discussed in a sentence and not repeated in subsequent phrases but referred to using pronouns, this should be coded as an accounting information to the same category as the previous.	"The operating expense ratio was 22.5% of sales and better than our estimate of 22.8%. **The improvement** was due to leverage associated with sales growth." (JPM on GPC, 2011)

Table 43: Acronyms, meaning and related categories

During the coding process, it was observed that analysts use several acronyms for accounting information.
This table presents the list of ones observed, their meaning and the category they belong[45].

Acronym	Keyword	Category
A/R	Accounts receivable	Assets
ARC	Average Revenue per Carload	Revenue
ARPU	Average Revenue Per Unit/User	Revenue
ATOI	After Tax Operating Income	Earnings
CFFO	Cash flow from operations	Cash Flow
CFO	Cash flow from operations	Cash Flow
CFPS	Cash flow per share	Cash Flow
COGS	Cost of goods sold	Expenses
Comps	Comparable same store sales	Revenue
DACF	Debt adjusted cash flow	Cash Flow
DD&A	Depreciation, Depletion & Amortization.	Expenses
DOI	Days of inventory	Activity Ratio
DSO	Days Sales Outstanding	Activity Ratio
DTA	Deferred Tax Asset	Asset
EV/OI	Enterprise value / operating income	Valuation Models
FCF	Free Cash Flow	Cash Flow
FTM	Forward or future twelve months	Future-Oriented
G&A	General and Administrative Expenses	Expenses
GM	Gross margin	Profitability Ratios
GOS	Gain on Sale	Other Financial Performance
GPM	Gross profit margins or Gross processing margins.	Profitability Ratios
ID/ID sales	(Identical) same store sales	Revenue
LOE	Leasehold operating expenses	Expenses
ND/EBITDA	Net Debt/EBITDA	Solvency
NG EPS	Non GAAP EPS	Earnings
NOL	Net operating loss	Earnings
NOPAT	Net operating profit after tax.	Earnings
NTM	Next twelve months	Future-Oriented
OCF	Operating Cash Flow	Cash Flow
OCI	Other comprehensive income	Earnings

[45] The meaning of the acronyms were derived either from the report itself (e.g. same information is conveyed using the acronym in one paragraph and the full word in another) or reference to other financial sources, e.g. Investopedia. For each acronym, there could be several words associated, e.g. GM could be used for Gross Margins or General Manager and CFO could be used for Cash Flow from Operations or Chief Financial Officer. So in each case in which these acronyms are observed, the context and surrounding information were considered before coding.

Acronym	Keyword	Category
OIBA	Operating income before amortisation	Earnings
OIBDA	Operating income before depreciation and amortisation	Earnings
OM	Operating Margins	Profitability Ratios
OM	Operating Margin	Profitability Ratios
OPBDA	Operating Profit Before Depreciation and Amortization	Earnings
OPEB	other post-employment benefits	Other Financial Performance
OPMg	Operating Margins	Profitability Ratios
OpMgns	Operating Margins	Profitability Ratios
OR	Operating Ratio	Profitability Ratios
PER	Price-Earnings Ratio	Valuation Models
PTOI	Pre-Tax Operating Income	Earnings
REVPAR	Revenue per available room	Revenue
Revs	Revenue	Revenue
RNOA	return on net operating asset	Profitability Ratios
RPU	Revenue Per Unit	Revenue
SSS	Same-store sales	Revenue
TBV	Tangible book value	Book Value
TEV/Sales	Total enterprise value/sales	Valuation Models
TTM	Trailing twelve months	Past-Oriented
UFCF	Unlevered Free Cash Flow	Cash flow

Appendix C: An illustrative example of computing Cohen's kappa

This appendix shows the computation of the Kappa coefficient which is used to estimate the inter-coder and intra-coder reliability in this study as presented in Chapter 5 of the Thesis.

		First coder		Total
		Non-Accounting	**Accounting**	
Second Coder	**Non-Accounting**	86	3	**89**
	Accounting	1	106	**107**
Total		**87**	**109**	**196**

A	Proportion of Accounting-First coder	109/196	**0.56**
B	Proportion of Accounting-Second coder	107/196	**0.55**
C	Proportion of Non-Accounting-First coder	87/196	**0.44**
D	Proportion of Non-Accounting-Second coder	89/196	**0.45**

Percentage Agreement		192/196	**98%**
Percentage expected agreement		(A*B)+ (C*D)	**0.51**
Kappa	(Percent Agreement - Expected agreement)/(1-expected agreement)		**96%**

Appendix D: Coded and misclassified words per report

This appendix provides examples of coded and misclassified words from the sensitivity analysis discussed in Chapter 9, extending Table 29. Table 29 contains the number of coded and misclassified words, this appendix shows the actual words coded.

The first table shows the coded and misclassified words for two reports and for the positive category of each wordlist. The second table shows the coded and misclassified words for other two reports for the negative category of each of the lists.

Panel A: SLB_Jeffries

SLB_Jeffries					
Dict_Pos		**LM_Pos**		**H_Pos**	
Words Captured	**Misclassified**	**Words Captured**	**Misclassified**	**Words Captured**	Misclassified
better (4)		achieved (1)		above (4)	
fair (1)	fair (1)	benefit (2)		achieved (1)	
good (1)		benefited (2)		beat (7)	
growth (5)		better (4)		better (4)	
healthy (3)		bolstered (5)		exceed (4)	
improvement (11)		boosted (2)		expansion (1)	
just (1)	just (1)	despite (1)		good (1)	
like (1)	like (1)	easily (1)		grew (6)	
noted (3)	noted (3)	efficiency (3)		grows (1)	
outstanding (1)		favorable (2)		growth (5)	
positive (8)		gain (11)	gain (1)	high (3)	
progress (1)		gained (1)		higher (16)	higher (2)
promise (1)	promise (1)	good (1)		improved (3)	
prosper (1)		improved (3)		improvement (14)	
security (1)	security (1)	improvement (11)		improvements (3)	
strong (9)		improvements (3)		improving (1)	
stronger (4)		improving (1)		increase (6)	
surprise (1)	surprise (1)	leading (2)		increased (14)	
worth (2)	worth (2)	outperforms (1)		increases (5)	
		Positive (8)		increasing (3)	
		profitability (2)		leading (2)	
		progress (1)		more (5)	more (3)
		rebound (1)		positive (9)	
		strength (4)		progress (1)	
		strong (9)		record (1)	
		stronger (4)		rise (4)	
		surpassed (4)		rising (1)	
				rose (6)	
				strength (4)	
				strong (9)	
				stronger (4)	
				up (10)	up (1)
59	11	90	1	155	6

Panel B: Goog_Indigo

Google_Indigo					
Dict_Pos		**LM_Pos**		**H_Pos**	
Words Captured	**Misclassified**	**Words Captured**	**Misclassified**	**Words Captured**	**Misclassified**
accuracy (1)		advantages (2)		better (1)	
attracting (1)		alliance (1)	alliance (1)	excellent (2)	
attractive (1)		attractive (1)		expanded (1)	
attracts (1)		better (1)		expanding (1)	
better (1)		compliment (1)		greater (1)	greater (1)
clear (1)		despite (1)		grow (2)	
competent (1)		effective (1)		grown (1)	
encourage (1)		enabled (1)		grows (3)	
excellent (2)		excellent (2)		growth (15)	
greater (1)	greater (1)	exceptional (1)		high (5)	
growth (15)	like (1)	exceptionally (1)		Improving (2)	
like (1)		gaining (1)		increase (1)	
promise (1)		greater (1)	greater (1)	Increased (2)	increased (3)
security (1)	security (1)	improving (2)		increasing (3)	
strong (5)		innovation (1)		largest (2)	
successful (1)		opportunities (1)		more (5)	more (2)
trust (1)	trust (1)	rebounded (1)		opportunities (1)	
		strengths (1)		record (2)	
		strong (5)		strengths (1)	
		successful (1)		strong (5)	
				successful (1)	
				up (2)	up (1)
36	4	27	2	59	8

Panel A: STT_Morningstar

STT_Morningstar					
Dict_Neg		**LM_Neg**		**H_Neg**	
Words Captured	**Misclassified**	**Words Captured**	**Misclassified**	**Words Captured**	**Misclassified**
adverse (1)		adverse (1)		below (1)	
crisis (2)		against (1)	against (1)	decline (2)	
difficulties (1)		argued (1)	argued (1)	declined (2)	declined (1)
disaster (2)		concerned (1)		downturn (1)	
disruption (1)		crisis (2)		drop (3)	
exploit (1)		cut (2)		less (2)	less (2)
losses (3)		decline (2)		low (2)	
negative (2)		declined (2)	declined (1)	lower (2)	lower (2)
problem (2)		detrimental (1)		negative (2)	
risk (11)		difficulties (1)		risk (11)	
stop (1)	stop (1)	disaster (2)		shrinking (2)	
stress (2)		disproportionate (1)	disproportionate (1)	smaller (1)	smaller (1)
suffer (1)		disruption (1)		threat (1)	
threat (1)		distract (1)		under (9)	under (9)
without (2)	without (2)	downturn (1)		worse (2)	
worse (2)		exploit (1)	exploit (1)		
		exposed (1)			
		exposing (1)			
		force (1)			
		losses (3)			
		mismanagement (1)			
		mistakes (1)			
		negative (2)			
		prevents (1)	prevents (1)		
		problem (2)			
		prolonged (1)			
		reckless (2)			
		slower (1)			
		stress (2)			
		suffer (1)			
		threat (1)			
		unsustainable (1)			
		unwanted (1)			
		worse (2)			
		wrongly (1)			
36	3	47	6	43	15

Panel B: WMT_Wells

WMT_Wells					
Dict_Neg		**LM_Neg**		**H_Neg**	
Words Captured	**Misclassified**	**Words Captured**	**Misclassified**	**Words Captured**	**Misclassified**
failure (1)		absence (1)		below (6)	none
feared (2)		against (1)	against (1)	decline (1)	
gross (6)	gross (6)	decline (1)		decrease (1)	
negative (5)		delays (1)		down (2)	
profit (2)	profit (2)	deliberate (1)		failure (1)	
risks (1)		distraction (1)		fell (1)	
weak (1)		failure (1)		low (1)	
weaker (1)		negative (5)		lower (12)	
without (2)	without (2)	negatively (1)		negative (5)	
worse (2)		shortfall (2)		risks (1)	
		slower (1)		weak (1)	
		weak (1)		worse (2)	
		weaker (1)			
		worse (1)			
23	10	20	1	34	0

Appendix E: Variable definition

Variables	Definition
ACC	Number of words in text units containing accounting related information divided by total word count
ACT	Number of words in text units containing activity ratios related information divided by total word count
ASS	Number of words in text units containing asset related information divided by total word count
Bog	Bog index
CF	Number of words in text units containing cash flow related information divided by total word count
CFA	Indicator variable which equals one if the report is issued by atleast a CFA qualified analyst and zero otherwise
EARN	Number of words in text units containing earnings related information divided by total word count
EQ	Number of words in text units containing equity related information divided by total word count
EXP	Number of words in text units containing expense related information divided by total word count
Financial Performance	Percentage change in EPS from prior year
FPER	Number of words in text units containing financial performance related information divided by total word count
FPOS	Number of words in text units containing financial position related information divided by total word count
Growth	Cumulative annual growth rate in sales per share over the previous five years (2006-2010)
IB	Indicator variable which equals one if the report is issued by an IB-analyst and zero if issued by and IND-analyst
Intangibles	Proportion of intangible assets over total assets for the fiscal year end 2010
Leverage	Long term debt to total asset ratio
LIAB	Number of words in text units containing liability related information divided by total word count
LIQ	Number of words in text units containing liquidity ratios related information divided by total word count
LTC	Indicator variable that equals one if the company is a low-tech company and zero otherwise.
MTB	Ratio of market value of equity to book value of equity for the fiscal year end 2010
NSA	Number of words in text units containing non-specific accounting related information divided by total word count
OACC	Number of words in text units containing other accouting information related information divided by total word count
OFPERF	Number of words in text units containing other financial performance related information divided by total word count
OFPOS	Number of words in text units containing other financial position related information divided by total word count
POS	Indicator variable which equals one if the recommendation contained in the report is positive and zero otherwise
PRATIO	Number of words in text units containing profitability ratios related information divided by total word count
REV	Number of words in text units containing revenue related information divided by total word count
Risk	Standard deviation of EBIT over five years (2006-2010), scaled by average assets over the same time period
Size	Natural logarithm of market value for the fiscal year end 2010
SOL	Number of words in text units containing solvency ratios related information divided by total word count
Tone	Number of positive words less number of negative words divided by the total word count (%).
VM	Number of words in text units containing valuation model related information divided by total word count

List of References

Abdolmohammadi, M., Simnett, R., Thibodeau, J. C., & Wright, A. M. (2006). Sell-side analysts' reports and the current external reporting model. *Accounting Horizons*, *20*(4), 375–389.

Abeysekera, I. (2006). The project of intellectual capital disclosure: researching the research. *Journal of Intellectual Capital*, *7*(1), 61–77.

Abhayawansa, S. A. (2010). *Sell-side analysts' use and communication of intellectual capital information*. PhD Thesis. University of Sydney.

Abhayawansa, S. (2011). A methodology for investigating intellectual capital information in analyst reports. *Journal of Intellectual Capital*, *12*(3), 446–476.

Abhayawansa, S. A., & Guthrie, J. (2012). Intellectual capital information and stock recommendations : Impression management? *Journal of Intellectual Capital*, *13*(3), 398–415.

Abhayawansa, S., Aleksanyan, M., & Bahtsevanoglou, J. (2015). The use of intellectual capital information by sell-side analysts in company valuation. *Accounting and Business Research*, *45*(3), 279–306.

Abraham, S., & Cox, P. (2007). Analysing the determinants of narrative risk information in UK FTSE 100 annual reports. *The British Accounting Review*, *39*(3), 227–248.

Aerts, W. (2005). Picking up the pieces: impression management in the retrospective attributional framing of accounting outcomes. *Accounting, Organizations and Society*, *30*(6), 493–517.

Agrawal, A., Chadha, S., & Chen, M. A. (2006). Who is afraid of Reg FD? The behavior and performance of sell-side analysts following the SEC's fair disclosure rules. *The Journal of Business*, *79*(6), 2811–2834.

Agrawal, A., & Chen, M. A. (2012). Analyst conflicts and research quality. *Quarterly Journal of Finance*, *2*(2), 1–40.

AICPA. (1994). *Improving business reporting—A customer focus: meeting the information needs of investors and creditors; comprehensive report of the special committee on financial reporting (The Jenkins Report)*. New York.

Akerlof, G. A. (1970). The market for "lemons": quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500.

Amir, E., & Lev, B. (1996). Value-relevance of non-financial information: The wireless communications industry. *Journal of Accounting and Economics*, *22*(1), 3–30.

Anderson, M. J. (1988). A comparative analysis of information search and evaluation behaviour of professional and non-professional financial analysts. *Accounting, Organisations and Society*, *13*(5), 431–446.

Arnold, J., & Moizer, P. (1984). A survey of the methods used by UK investment analysts to appraise investments in ordinary shares. *Accounting and Business Research*, *14*(55), 195–207.

Asquith, P., Mikhail, M. B., & Au, A. S. (2005). Information content of equity analyst reports. *Journal of Financial Economics*, *75*(2), 245–282.

Bailey, W., Li, H., Mao, C. X., & Zhong, R. (2003). Regulation fair disclosure and earnings information: market, analyst, and corporate responses. *Journal of Finance*, *58*(6), 2487–2514.

Ball, R., & Shivakumar, L. (2008). How much new information is there in earnings? *Journal of Accounting Research*, *46*(5), 975–1016.

Barber, B. M., Lehavy, R., & Trueman, B. (2007). Comparing the stock recommendation performance of investment banks and independent research firms. *Journal of Financial Economics*, *85*(2), 490–517.

Barker, R. (1998). The market for information—evidence from finance directors, analysts and fund managers. *Accounting and Business Research*, *29*(1), 3–20.

Barker, R. (1999). The role of dividends in valuation models used by analysts and fund managers. *European Accounting Review*, *8*(2), 195–218.

Barker, R. G. (2000). FRS3 and analysts' use of earnings. *Accounting and Business Research*, *30*(2), 95–109.

Barker, R. (2001). *Institutional investors, accounting information and the ASB*. *The Institute of Chartered Accountants of Scotland*.

Barker, R., Hendry, J., Roberts, J., & Sanderson, P. (2012). Can company-fund manager meetings convey informational benefits? Exploring the rationalisation of equity investment decision making by UK fund managers. *Accounting, Organizations and Society*, *37*(4), 207–222.

Barker, R., & Imam, S. (2008). Analysts' perceptions of earnings quality. *Accounting and Business Research*, *38*(4), 313–329.

Barniv, R., Hope, O., Myring, M. J., & Thomas, W. B. (2009). Do analysts practice what they preach and should investors listen? Effects of recent regulations. *The Accounting Review*, *84*(4), 1015–1039.

Barth, M. E., Beaver, W. H., & Landsman, W. R. (1998). Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting and Economics*, *25*(1), 1–34.

Barth, M. E., Beaver, W. H., & Landsman, W. R. (2001). The relevance of the value relevance literature for financial accounting standard setting: another view. *Journal of Accounting and Economics*, *31*(1), 77–104.

Barth, M. E., Landsman, W. R., & Lang, M. H. (2008). International accounting standards and accounting quality. *Journal of Accounting Research*, *46*(3), 467–498.

Barton, J., Hansen, T. B., & Pownall, G. (2010). Which performance measures do investors around the world value the most-and why? *Accounting Review*, *85*(3), 753–789.

Beattie, V. (2014). Accounting narratives and the narrative turn in accounting research: Issues, theory, methodology, methods and a research framework. *British Accounting Review*, *46*(2), 111–134.

Beattie, V., Dhanani, A., & Jones, M. J. (2008). Investigating presentational change in UK annual reports: A longitudinal perspective. *Journal of Business Communication*, *45*(2), 181–222.

Beattie, V., McInnes, B., & Fearnley, S. (2004). A methodology for analysing and evaluating narratives in annual reports: a comprehensive descriptive profile and metrics for disclosure quality attributes. *Accounting Forum*, *28*(3), 205–236.

Beattie, V., & Thomson, S. J. (2007). Lifting the lid on the use of content analysis to investigate intellectual capital disclosures. *Accounting Forum*, *31*(2), 129–163.

Beisland, L. A. (2009). A review of the value relevance literature. *The Open Business Journal*, *2*(1), 7–27.

Beunza, D., & Garud, R. (2007). Calculators, lemmings or frame-makers? The intermediary role of securities analysts. *Sociological Review*, *55*(S.2), 13–39.

Beyer, A., Cohen, D. a., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: review of the recent literature. *Journal of Accounting and Economics*, *50*(2-3), 296–343.

Biggs, S. F. (1984). Financial analysts' information search in the assessment of corporate earning power. *Accounting, Organisations and Society*, *9*(3), 313–323.

Block, S. (1999). A study of financial analysts: practice and theory. *Financial Analysts Journal*, *55*(4), 86–95.

Boni, L., & Womack, K. L. (2002). *Solving the sell-side research problem : insights from buy-side professionals*. *Unpublished Manuscript*.

Bonsall, S. B. V, Leone, A. J., & Miller, B. P. (2015). A plain English measure of financial reporting readability. *Working Paper*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2560644

Bouwman, M. J., Frishkoff, P., & Frishkoff, P. A. (1995). The relevance of GAAP-based information : A case study exploring some uses and limitations. *Accounting Horizons*, *9*(4), 22–47.

Bowen, R. M. (2008). Analyst coverage and the cost of raising equity capital : evidence from under-pricing of seasoned equity offerings. *Contemporary Accounting Research*, *25*(3).

Boyatzis, R. E. (1998). *Transforming qualitative information : thematic analysis and code development*. London: Sage Publications.

Bradshaw, M. T. (2002). The use of target prices to justify sell-side analysts' stock recommendations. *Accounting Horizons*, *16*(1), 27–41.

Bradshaw, M. T. (2004). How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review*, *79*(1), 25–50.

Bradshaw, M. T. (2011). *Analysts' forecasts: what do we know after decades of work? Working Paper*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1880339

Brennan, N. M., Guillamon-Saorin, E., & Pierce, A. (2009). Impression management: Developing and illustrating a scheme of analysis for narrative disclosures – a methodological note. *Accounting, Auditing & Accountability Journal*, *22*(5), 789–832.

Brennan, N. M., & Merkl-Davies, D. M. (2013). Accounting narratives and impression management. In *The Routledge Companion to Communication in Accounting* (pp. 109–132). London: Routledge.

Breton, G., & Taffler, R. J. (2001). Accounting information and analyst stock recommendation decisions: a content analysis approach. *Accounting and Business Research*, *31*(2), 91–101.

Bricker, R., Previts, G., Robinson, T., & Young, S. (1995). Financial analyst assessment of a company's earnings quality. *Journal of Accounting, Auditing & Finance*, *10*(3), 541–554.

Brown, D. L. (1993). Earnings forecasting research : its implications for capital markets research. *International Journal of Forecasting*, *9*(3), 295–320.

Brown, L. D., Call, A. C., Clement, M. B., & Sharp, N. Y. (2015). Inside the "black box" of sell-side financial analysts. *Journal of Accounting Research*, *53*(1), 1–47.

Bryan, D. M., & Tiras, S. L. (2007). The Influence of forecast dispersion on the incremental explanatory power of earnings, book value, and analyst forecasts on market prices. *The Accounting Review*, *82*(3), 651–677.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, *6*(3), 203–242.

Burgstahler, D. C., & Dichev, I. D. (1997). Earnings, adaptation and equity value. *The Accounting Review*, *72*(2), 187–215.

Bushee, B. J., Gow, I. D., & Taylor, D. J. (2013). *Linguistic complexity in firm disclosures : Obfuscation or information ? Working Paper*. Retrieved from http://ssrn.com/abstract=2375424.

Bushman, R. M., & Smith, A. J. (2001). Financial accounting information and corporate governance. *Journal of Accounting & Economics*, *32*(1), 237–333.

Cameron, C. A., & Miller, D. L. (2013). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317–372.

Campbell, D., & Slack, R. (2008). *Narrative reporting: analysts' perceptions of its value and relevance*. *Association of Chartered Certified Accountants*.

Carney, T. F. (1972). *Content analysis : a technique for systematic inference from communications*. London: Batsford.

Cascino, S., Clatworthy, M., Osma, B. G., Gassen, J., Imam, S., & Jeanjean, T. (2013). The use of information by capital providers: Academic literature review. *Institute of Chartered Accountants of Scotland and European Financial Reporting Advisory Group. Available via www. efrag. org*.

Chandra, U., & Ro, B. T. (2008). The Role of revenue in firm valuation. *Accounting Horizons*, *22*(2), 199–222.

Charitou, A., Lambertides, N., & Trigeorgis, L. (2011). Distress risk, growth and earnings quality. *Abacus*, *47*(2), 158–181.

Chen, C., & Chen, P. F. (2009). NASD Rule 2711 and changes in analysts' independence in making stock recommendations. *The Accounting Review*, *84*(4), 1041–1071.

Chen, S., & Matsumoto, D. a. (2006). Favourable versus unfavourable recommendations: The impact on analyst access to management-provided information. *Journal of Accounting Research*, *44*(4), 657–689.

Chen, T., Harford, J., & Lin, C. (2012). *Do financial analysts play a monitoring role? Evidence from natural experiments*. *Working paper*. Retrieved from http://www.business.illinois.edu/Finance/papers/2013/Harford.pdf

Chen, X., Cheng, Q., & Lo, K. (2010). On the relationship between analyst reports and corporate disclosures: Exploring the roles of information discovery and interpretation. *Journal of Accounting and Economics*, *49*(3), 206–226.

Cho, C. H., Roberts, R. W., & Patten, D. M. (2010). The language of US corporate environmental disclosure. *Accounting, Organizations and Society*, *35*(4), 431–443.

Chugh, L., & Meador, J. (1984). The stock valuation process: The analysts' view. *Financial Analysts Journal*, *40*(6), 41–48.

Clarke, J. E., Khorana, A., Patel, A., & Rau, P. R. (2011). Independents' day? Analyst behaviour surrounding the Global Settlement. *Annals of Finance*, *7*(4), 529–547.

Clatworthy, M. A., & Jones, M. J. (2006). Differential patterns of textual characteristics and company performance in the chairman's statement. *Accounting, Auditing & Accountability Journal*, *19*(4), 493–511.

Clatworthy, M., & Jones, M. J. (2001). The effect of thematic structure on the variability of annual report readability. *Accounting, Auditing & Accountability Journal*, *14*(3), 311–326.

Clatworthy, M., & Jones, M. J. (2003). Financial reporting of good news and bad news : evidence from accounting narratives. *Accounting and Business Research*, *33*(3), 171–185.

Coffee, J. C. (2002). Understanding Enron : It's about the gatekeepers, stupid. *The Business Lawyer*, 1403–1420.

Coleman, I., & Eccles, R. (1997). *Pursuing value: reporting gaps in the United Kingdom*.

Collins, D. W., & Kothari, S. P. (1989). An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics*, *11*(2), 143–181.

Collins, D. W., Maydew, E. L., & Weiss, I. S. (1997). Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics*, *24*(1), 39–67.

Coram, P. J., Mock, T. J., & Monroe, G. S. (2011). Financial analysts' evaluation of enhanced disclosure of non-financial performance indicators. *The British Accounting Review*, *43*(2), 87–101.

Courtis, J. K. (1998). Annual report readability variability: tests of obfuscation hypothesis. *Accounting, Auditing & Accountability Journal*, *11*(4), 459–472.

Courtis, J. K. (2004). Corporate report obfuscation : artefact or phenomenon ? *The British Accounting Review*, *36*(3), 291–312.

Cowen, A., Groysberg, B., & Healy, P. (2006). Which types of analyst firms are more optimistic? *Journal of Accounting and Economics*, *41*(1), 119–146.

Cram, D. P., Karan, V., & Stuart, I. (2009). Three threats to validity of choice-based and matched-sample studies in accounting research. *Contemporary Accounting Research*, *26*(2), 477–516.

Darrough, M., & Ye, J. (2007). Valuation of loss firms in a knowledge-based economy. *Review of Accounting Studies*, *12*(1), 61–93.

Das, S., Levine, C. B., & Sivaramakrishnan, K. (1998). Earnings predictability and bias in analysts' earnings' forecasts. *The Accounting Review*, *73*(2), 277–294.

De Franco, G., Hope, O. K., Vyas, D., & Zhou, Y. (2015). Analyst report readability. *Contemporary Accounting Research*, *32*(1), 76–104.

De Franco, G., & Zhou, Y. (2009). The performance of analysts with a CFA designation: The role of human-capital and signalling theories. *The Accounting Review, 84*(2), 383–404.

Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics, 50*(2), 344–401.

Demers, E., & Vega, C. (2011). *Linguistic tone in earnings press releases: News or noise*. *Working paper*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1152326

Demirakos, E. G., Strong, N. C., & Walker, M. (2004). What valuation models do analysts use? *Accounting Horizons*, *18*(4), 221–240.

Der Hovanesian, M., & Borrus, A. (2005). Can the street make research pay? Retrieved May 6, 2015, from http://www.bloomberg.com/bw/stories/2005-01-30/can-the-street-make-research-pay

Dichev, I. D., & Tang, V. W. (2009). Earnings volatility and earnings predictability. *Journal of Accounting and Economics*, *47*(1), 160–181.

Doukas, J. A., Kim, C., & Pantzalis, C. (2000). Security analysis, agency costs, and company characteristics. *Financial Analysts Journal*, *56*(6), 54–63.

Dugar, A., & Nathan, S. (1995). The effect of investment banking relationships on financial analysts' earnings' forecasts and investment recommendations. *Contemporary Accounting Research*, *12*(1), 131–160.

Eames, M., Glover, S. M., & Kennedy, J. (2002). The association between trading recommendations and broker-analysts' earnings forecasts. *Journal of Accounting Research, 40*(1), 85–104.

Eames, M. J., & Glover, S. M. (2003). Earnings predictability and the direction of analysts' earnings forecast errors. *The Accounting Review*, *78*(3), 707–724.

Easton, P. D. (1999). Security returns and the value relevance of accounting data. *Accounting Horizons, 13*(4), 399–412.

Epstein, M. J., & Palepu, K. G. (1999). What financial analysts want? *Strategic Finance, 80*(10), 1–7.

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, *15*(4), 915–953.

Feng, M., & Mcvay, S. (2010). Analysts' incentives to overweight management guidance when revising their short-term earnings forecasts. *The Accounting Review, 85*(5), 1617–1646.

Fernandez, F. (2001). The roles and responsibilities of securities analysts. *SIFMA Research Report*, *2*(7), 3–10.

Financial Reporting Council. (2012). *Cutting clutter: Combatting clutter in annual reports*. Retrieved from www.frc.org.uk

Firth, M., Lin, C., Liu, P., & Xuan, Y. (2013). The client is king: do mutual fund relationships bias analyst recommendations? *Journal of Accounting Research*, *51*(1), 165–200.

Fisch, J. E., & Sale, H. A. (2002). Securities Analyst as Agent: Rethinking the Regulation of Analysts. *The Iowa Law Review, 88*, 1035–1098.

Flöstrand, P. (2006). The sell side – observations on intellectual capital indicators. *Journal of Intellectual Capital*, *7*(4), 457–473.

Flöstrand, P., & Ström, N. (2006). The valuation relevance of non-financial information. *Management Research News*, *29*(9), 580–597.

Fogarty, T. J., & Rogers, R. K. (2005). Financial analysts' reports: an extended institutional theory evaluation. *Accounting, Organizations and Society*, *30*(4), 331–356.

Francis, J., Nanda, D., & Wang, X. (2006). Re-examining the effects of regulation fair disclosure using foreign listed firms to control for concurrent shocks. *Journal of Accounting and Economics*, *41*(3), 271–292.

Francis, J., & Philbrick, D. (1993). Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research*, *31*(2), 216–230.

Francis, J., & Schipper, K. (1999). Have financial statements lost their relevance? *Journal of Accounting Research*, *37*(2), 319–352.

Francis, J., Willis, R. H., & Philbrick, D. R. (2004). *Security analyst independence*. *Chartered Financial Analysts Institute*.

Frank, K. (2002). The effect of growth on the value relevance of accounting data. *Journal of Business Research*, *55*(1), 69–78.

Frankel, R.; Kothari, S.P.; Weber, J. (2006). Determinants of the informativeness of analyst research. *Journal of Accounting and Economics, 41*(1), 29–54.

García-Meca, E., & Martínez, I. (2007). The use of intellectual capital information in investment decisions: An empirical study using analyst reports. *The International Journal of Accounting*, *42*(1), 57–81.

Gassen, J. (2008). *Are stewardship and valuation usefulness compatible or alternative objectives of financial accounting? Working paper*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1095215

Gassen, J. (2014). Causal inference in empirical archival financial accounting research. *Accounting, Organizations and Society*, *39*(7), 535–544.

General Accounting Office, U. S. (1989). *Content analysis : a methodology for structuring and analysing written material*.

Gertsen, F. H. M. (2009). *Riding a tiger without being eaten: How companies and analysts tame financial restatements and influence corporate reputation*. *PhD Thesis*. Erasmus University Rotterdam.

Ghosh, A. Al, & Moon, D. (2010). Corporate debt financing and earnings quality. *Journal of Business Finance & Accounting*, *37*(5-6), 538–559.

Glaum, M., & Friedrich, N. (2006). After the "bubble": valuation of telecommunications companies by financial analysts. *Journal of International Financial Management and Accounting*, *17*(2), 160–174.

Gordon, R. A. (1996). Impact of ingratiation on judgments and evaluations : A meta-analytic investigation. *Journal of Personality and Social Psychology*, *71*(1), 54–70.

Govindarajan, V. (1980). The objectives of financial statements: An empirical study of the use of cash flow and earnings by security analysts. *Accounting, Organizations and Society*, *5*(4), 383–392.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, *40*(1-3), 3–73.

Gray, R., Kouhy, R., & Lavers, S. (1995). Corporate social and environmental reporting: a review of the literature and a longitudinal study of UK disclosure. *Accounting, Auditing & Accountability Journal*, *8*(2), 47–77.

Green, T. C., Jame, R., Markov, S., & Subasi, M. (2014). Broker-hosted investor conferences. *Journal of Accounting and Economics*, *58*(1), 142–166.

Groysberg, B. (2013). *Wall Street Research: Past, Present, and Future*. Stanford University Press.

Groysberg, B., Healy, P. M., & Maber, D. A. (2011). What drives sell-side analyst compensation at high-status investment banks? *Journal of Accounting Research*, *49*(4), 969–1000.

Gu, Z., & Xue, J. (2008). The superiority and disciplining role of independent analysts. *Journal of Accounting and Economics*, *45*(2-3), 289–316.

Guan, Y., Lu, H., & Wong, M. H. F. (2012). Conflict-of-interest reforms and investment bank analysts' research biases. *Journal of Accounting, Auditing & Finance*, *27*(4), 443–470.

Guthrie, J., Petty, K., & Yongvanich, R. F. (2004). Using content analysis as a research method to inquire into intellectual capital reporting. *Journal of Intellectual Capital*, *5*(2), 282–293.

Hart, R. P. (2001). Redeveloping Diction: theoretical considerations. In *Progress in communication sciences* (pp. 43–60). London: Ablex Publishing.

Hassan, E. (2006). Recall Bias can be a threat to retrospective and prospective research designs. *The Internet Journal of Epidemiology*, *3*(2), 1–7.

Hayn, C. (1995). The information content of losses. *Journal of Accounting and Economics*, *20*(2), 125–153.

Hayward, M. L. A., & Boeker, W. (1998). Power and conflicts of interest in professional firms : evidence from investment banking. *Administrative Science Quarterly*, *43*(1), 1–22.

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, *31*(1), 405–440.

Heflin, F., Subramanyam, K. R., & Zhang, Y. (2003). Regulation FD financial information environment : early evidence. *The Accounting Review*, *78*(1), 1–37.

Henry, E. (2008). Are Investors Influenced By How Earnings Press Releases Are Written? *Journal of Business Communication*, *45*(4), 363–407.

Holland, J. (2004). *Corporate intangibles, value relevance and disclosure content*. *Institute of Chartered Accountants of Scotland*.

Holland, J. (2006). Fund management, intellectual capital, intangibles and private disclosure. *Managerial Finance*, *32*(4), 277–316.

Holland, J., & Johanson, U. (2003). Value- relevant information on corporate intangibles – creation, use, and barriers in capital markets – "between a rock and a hard place." *Journal of Intellectual Capital*, *4*(4), 465–486.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. London: Addison-Wesley.

Holthausen, R. W., & Watts, R. L. (2001). The relevance of the value-relevance literature for financial accounting standard setting. *Journal of Accounting and Economics*, *31*(1), 3–75.

Hong, H., & Kubik, J. D. (2003). Analyzing the analysts : Career concerns and biased earnings forecasts. *The Journal of Finance*, *58*(1), 313–351.

Hong, H., Lim, T., & Stein, J. C. (2000). Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance*, *55*(1), 265–295.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Third edit). Wiley.

Huang, A., Zang, A., & Zheng, R. (2014). Evidence on the informativeness of text in analyst reports. *The Accounting Review*, *89*(6), 2151–2180.

Imam, S. (2011). *What do analysts do ? Evidence from UK sell-side and buy-side analysts*. *Working Paper*.

Imam, S., Barker, R., & Clubb, C. (2008). The use of valuation models by UK investment analysts. *European Accounting Review*, *17*(3), 503–535.

Irani, R. M., & Oesch, D. (2013). Monitoring and corporate disclosure: Evidence from a natural experiment. *Journal of Financial Economics*, *109*(2), 398–418.

Irvine, P. J. (2003). The incremental impact of analyst initiation of coverage. *Journal of Corporate Finance*, *9*, 431–451.

Jacob, J., Rock, S., & Weber, D. P. (2008). Do non-investment bank analysts make better earnings forecasts? *Journal of Accounting, Auditing & Finance*, *23*(1), 23–61.

Jegadeesh, N., Kim, J., Krische, S. D., & Lee, C. M. C. (2004). Analyzing the analysts : when do recommendations add value ? *The Journal of Finance*, *59*(3), 1083–1124.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm : Managerial behavior , agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305–360.

Johnston, R., Markov, S., & Ramnath, S. (2009). Sell-side debt analysts. *Journal of Accounting and Economics*, *47*(1), 91–107.

Jones, E. E. (1964). *Ingratiation: A social psychological analysis*. New York: Appleton-Century-Crofts.

Jones, M. J. (1997). Critical appraisal of the cloze procedure's use in the accounting domain. *Accounting, Auditing & Accountability Journal*, *10*(1), 105–128.

Jones, M. J., & Shoemaker, P. A. (1994). Accounting narratives: a review of empirical studies of content and readability. *Journal of Accounting Literature*, *13*, 142–184.

Kacmar, K. M., Carlson, D. S., & Bratton, V. K. (2004). Situational and dispositional factors as antecedents of ingratiatory behaviours in organizational settings. *Journal of Vocational Behaviour*, *65*(2), 309–331.

Kadan, O., Madureira, L., Wang, R., & Zach, T. (2009). Conflicts of interest and stock recommendations: the effects of the Global Settlement and related regulations. *Review of Financial Studies*, *22*(10), 4189–4217.

Kang, Q. (2012). *CFA certification program and sell-side analysts CFA certification program and sell-side analysts*. *Working Paper*. Retrieved from SSRN 2137312

Ke, B., & Yu, Y. (2006). The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research*, *44*(5), 965–999.

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, *33*, 171–185.

Kennedy, P. (2008). *A guide to econometrics* (6th Edition). Blackwell Publishing.

Kirk, M., & Markov, S. (2015). *Come on over : Analyst / investor days as a disclosure medium*. *Working Paper*. Retrieved from http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2144972

Kolasinski, A. C., & Kothari, S. P. (2008). Investment banking and analyst objectivity: Evidence from analysts affiliated with mergers and acquisitions advisors. *Journal of Financial and Quantitative Analysis*, *43*(4), 817–842.

Kothari, S. P., Li, X., & Short, J. E. (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a study using content analysis. *The Accounting Review*, *84*(5), 1639–1670.

Krippendorff, K. (2004). *Content analysis : an introduction to its methodology*. London: Sage.

Kross, W. J., & Suk, I. (2012). Does Regulation FD work ? Evidence from analysts' reliance on public disclosure. *Journal of Accounting and Economics*, *53*(1), 225–248.

Kumar, K., & Beyerlein, M. (1991). Construction and validation of an instrument for measuring ingratiatory behaviours in organizational settings. *Journal of Applied Psychology*, *76*(5), 619–627.

Labhart, D. J. (2004). Securities analysts : why these gatekeepers abandoned their post. *Indiana Law Journal*, *79*(4), 1037–1061.

Lambert, R., Leuz, C., & Verrecchia, R. E. (2007). Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research*, *45*(2), 385–420.

Lang, M. H., & Lundholm, R. J. (1996). Corporate disclosure policy and analyst behaviour. *The Accounting Review*, *71*(4), 467–492.

Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, *86*(3), 1087–1115.

Lewis, N. R., Parker, L. D., Pound, G. D., & Sutcliffe, P. (1986). Accounting report readability: the use of readability techniques. *Accounting and Business Research*, *16*(63), 199–213.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, *45*, 221–247.

Li, F. (2010). The information content of forward-looking statements in corporate filings-a naïve Bayesian machine learning approach. *Journal of Accounting Research*, *48*(5), 1049–1102.

Libby, R., & Lewis, B. L. (1977). Human information processing research in accounting: the state of the art. *Accounting, Organizations and Society*, *2*(3), 245–268.

Lin, H., & McNichols, M. F. (1998). Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics*, *25*(1), 101–127.

Linsley, P. M., & Lawrence, M. J. (2007). Risk reporting by the largest UK companies: readability and lack of obfuscation. *Accounting, Auditing & Accountability Journal*, *20*(4), 620–627.

Linsley, P. M., & Shrives, P. J. (2006). Risk reporting: A study of risk disclosures in the annual reports of UK companies. *The British Accounting Review*, *38*(4), 387–404.

Livnat, J., & Zhang, Y. (2012). Information interpretation or information discovery: Which role of analysts do investors value more? *Review of Accounting Studies*, *17*(3), 612–641.

Ljungqvist, A., Marston, F., Starks, L. T., Wei, K. D., & Yan, H. (2007). Conflicts of interest in sell-side research and the moderating role of institutional investors. *Journal of Financial Economics*, *85*(2), 420–456.

Lo, K., & Lys, T. (2000). The Ohlson model : contribution to valuation theory, limitations, and empirical applications. *Journal of Accounting, Auditing & Finance*, *15*(3), 337–367.

Long, S. J. (1997). *Regression models for categorical and limited dependent variables*. London: Sage.

Loughran, T. I. M., & Mcdonald, B. (2011). When is a liability not a liability ? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *LXVI* (1), 35–65.

Loughran, T., & Mcdonald, B. (2014). Measuring readability in financial disclosures. *Journal of Finance*, *69*(4), 1643–1671.

Lui, D., Markov, S., & Tamayo, A. (2007). What makes a stock risky? Evidence from sell-side analysts' risk ratings. *Journal of Accounting Research*, *45*(3), 629–665.

Maber, D., Groysberg, B., & Healy, P. M. (2014). *The use of broker votes to reward brokerage firms' and their analysts' research activities*. *Working Paper*. Retrieved from http://ssrn.com/abstrract=2311152

Malmendier, U., & Shanthikumar, D. (2014). Do security analysts speak in two tongues? *Review of Financial Studies*, *27*(5), 1287–1322.

Marston, C. (2004). *A Survey of European Investor Relations*. *Institute of Chartered Accountants of Scotland*.

Matsumoto, K., Shivaswamy, M., & Hoban, J. P. (1995). Security analysts' views of the financiai ratios of manufacturers and retailers. *Financial Practice and Education*, *5*(2), 44–55.

Mayew, W. J. (2008). Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research*, *46*(3), 627–659.

McEwen, R. A., & Hunton, J. E. (1999). Is analyst forecast accuracy associated with accounting information use? *Accounting Horizons*, *13*(1), 1–16.

McInnes, B., Beattie, V., & Pierpoint, J. (2007). *Communication between management and stakeholders: A case study*.

Mehran, H., & Stulz, R. M. (2007). The economics of conflicts of interest in financial institutions. *Journal of Financial Economics*, *85*(2), 267–296.

Merkl-Davies, D. M. (2007). *The obfuscation hypothesis re-examined : Analyzing impression management in corporate narrative report documents*. *PhD Thesis*. University of Wales.

Merkl-Davies, D. M., & Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives : incremental information or impression management ? *Journal of Accounting Literature*, *27*, 116–196.

Merkl-Davies, D. M., & Brennan, N. M. (2011). A conceptual framework of impression management : new insights from psychology, sociology and critical perspectives. *Accounting and Business Research*, *41*(5), 415–437.

Merkl-Davies, D. M., Brennan, N. M., & McLeay, S. J. (2011). Impression management and retrospective sense-making in corporate narratives: A social psychology perspective. *Accounting, Auditing & Accountability Journal*, *24*(3), 315–344.

Merkl-Davies, D. M., Brennan, N. M., & Vourvachis, P. (2014). Content analysis and discourse analysis in corporate narrative reporting research: a methodological guide. In *Crtical Perspectives in Accounting Conference*.

Michaely, R., & Womack, K. L. (1999). Conflict of interest and the credibility of underwriter analyst recommendations. *The Review of Financial Studies*, *12*(4), 653–686.

Miller, B. P. (2010). The Effects of reporting complexity on small and large investor trading. *The Accounting Review*, *85*(6), 2107–2143.

Milne, M., & Adler, R. (1999). Exploring the reliability of social and environmental disclosures content analysis. *Accounting, Auditing & Accountability Journal*, *12*(2), 237–256.

Moizer, P., & Arnold, J. (1984). Share appraisal by investment analysts—portfolio vs. non-portfolio managers. *Accounting and Business Research*, *14*(56), 341–348.

Mokoaleli-Mokoteli, T., Taffler, R. J., & Agarwal, V. (2009). Behavioural bias and conflicts of interest in analyst stock recommendations. *Journal of Business Finance & Accounting*, *36*(3-4), 384–418.

Morris, R. (1994). Computerized content analysis in management research: A demonstration of advantages & limitations. *Journal of Management*, *20*(4), 903–931.

Neuendorf, K. A. (2002). *The content analysis guidebook*. London: Sage Publications.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words : Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, *29*(5), 665–675.

Nielsen, C. (2007). A content analysis of analyst research: Healthcare through the eye of analysts. *Journal of Health Care Finance*, *34*(3), 66–90.

O'Brien, P. C., Mcnichols, M. F., & Hsiou-Wei, L. (2005). Analyst impartiality and investment banking relationships. *Journal of Accounting Research*, *43*(4), 623–650.

Ohlson, J. a. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research*, *11*(2), 661–687.

Orens, R., & Lybaert, N. (2007). Does the financial analysts' usage of non-financial information influence the analysts' forecast accuracy? Some evidence from the Belgian sell-side financial analyst. *The International Journal of Accounting*, *42*(3), 237–271.

Orens, R., & Lybaert, N. (2010). Determinants of sell-side financial analysts' use of non-financial information, *40*(1), 39–53.

Orpurt, S. F., & Zhang, Y. (2009). Do direct cash flow disclosures help predict future operating cash flows and earnings? *The Accounting Review*, *84*(3), 893–935.

Pandey, J., & Bohra, K. A. (1984). Ingratiation a s a function of organizational characteristics and supervisory styles. *International Review of Applied Psychology*, *33*(3), 381–394.

Patelli, L., & Pedrini, M. (2014). Is the optimism in CEO's letters to shareholders sincere? Impression management versus communicative action during the economic crisis. *Journal of Business Ethics*, *124*(1), 19–34.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: our words, our selves. *Annual Review of Psychology*, *54*(1), 547–577.

Pike, R., Meerjanssen, J., & Chadwick, L. (1993). The appraisal of ordinary shares by investment analysts in the UK and Germany. *Accounting and Business Research*, *23*(92), 489–499.

Previts, G. J., Bricker, R. J., Robinson, T. R., & Young, S. J. (1994). A content analysis of sell-side financial analyst company reports. *Accounting Horizons*, *8*(2), 55–70.

Ramnath, S., Rock, S., & Shane, P. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. *International Journal of Forecasting*, *24*(1), 34–75.

Rennekamp, K. (2012). *The influence of performance and reporting goals on managers' choice of reporting complexity in disclosures*. *Working Paper*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2137133

Richardson, S., Teoh, S. H., & Wysocki, P. D. (2004). The Walk-down to beatable analyst forecasts: the role of equity issuance and insider trading incentives. *Contemporary Accounting Research*, *21*(4), 885–924.

Robertson, J. C. (2010). The effects of ingratiation and client incentive on auditor judgment. *Behavioural Research in Accounting*, *22*(2), 69–86.

Rogers, J. L., Van Buskirk, A., & Zechman, S. L. C. (2011). Disclosure tone and shareholder litigation. *The Accounting Review*, *86*(6), 2155–2183.

Rogers, R. K., & Grant, J. (1997). Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis*, *3*(1), 17–30.

Roulstone, D. T. (2003). Analyst Following and Market Liquidity. *Contemporary Accounting Research*, *20*(3), 551–578.

Schipper, K. (1991). Analysts' Forecasts. *Accounting Horizons*, *5*(4), 105–121.

Schleicher, T., & Walker, M. (2010). Bias in the tone of forward looking narratives. *Accounting and Business Research*, *40*(4), 371–390.

Schlenker, B. R. (1980). *Impression management: The self-concept, social identity and interpersonal relations*. Monterey, California: Brooks/Cole Publishing Company.

Securities and Exchange Commission. (1998). *A plain English handbook: how to create clear SEC disclosure documents*.

Sirri, E. (2004). Investment banks, scope, and unavoidable conflicts of interest. *Economic Review*, *Quarter 4*, 23–35.

Smith, M., & Taffler, R. (2000). The chairman's statement-a content analysis of discretionary narrative disclosures. *Accounting, Auditing & Accountability Journal*, *13*(5), 624–647.

Soltes, E. (2014). Private interaction between firm management and sell-side analysts. *Journal of Accounting Research*, *52*(1), 245–272.

Steenkamp, N., & Northcott, D. (2007). Content analysis in accounting research: the practical challenges. *Australian Accounting Review*, *17*(3), 12–25.

Suozzo, P., Copper, S., Sutherland, G., & Deng, Z. (2001). *Valuation Multiples : A primer global equity. Global Equity Research*.

Swedberg, R. (2005). Conflicts of interests in the U.S. brokerage industry. In *The sociology of financial markets* (pp. 187–203). Oxford University Press.

Sydserff, R., & Weetman, P. (1999). A texture index for evaluating accounting narratives; An alternative to readability formulas. *Accounting, Auditing & Accountability Journal*, *12*(4), 459–488.

Sydserff, R., & Weetman, P. (2002). Developments in content analysis: a transitivity index and DICTION scores. *Accounting, Auditing & Accountability Journal*, *15*(4), 523–545.

Tedeschi, J. T., & Melburg, V. (1984). Impression management and influence in the organization. *Research in the Sociology of Organizations*, *3*, 31–58.

The Institute of Chartered Accountants in England and Wales. (2009). *Developments in new reporting models. Information for Better Markets Initiative*.

Twedt, B., & Rees, L. (2012). Reading between the lines: An empirical examination of qualitative attributes of financial analysts' reports. *Journal of Accounting and Public Policy*, *31*(1), 1–21.

Umiastowski, C. (2014). What to read (and what to ignore) in analyst reports. Retrieved July 1, 2015, from http://www.theglobeandmail.com/globe-investor/investment-ideas/strategy-lab/growth-investing/what-to-read-and-what-to-ignore-in-analyst-reports/article16353898/

Unerman, J. (2000). Methodological issues-reflections on quantification in corporate social reporting content analysis. *Accounting, Auditing & Accountability Journal*, *13*(5), 667–681.

Valentine, J. (2010). *Best Practices for Equity Research Analysts: Essentials for Buy-Side and Sell-Side Analysts* (First). New York: McGraw-Hill.

Vergoossen, R. (1993). The use and perceived importance of annual reports by investment analysts in the Netherlands. *European Accounting Review*, *2*(2), 219–244.

Vourvachis, P. (2007). On the use of content analysis (CA) in corporate social reporting (CSR): revisiting the debate on the units of analysis and the ways to define them. In *British Accounting Association Annual Conference 2007*. Egham, England. Retrieved from http://eprints.kingston.ac.uk/4129/

Wang, I. Y. (2007). Private earnings guidance and its implications for disclosure regulation. *The Accounting Review*, *82*(5), 1299–1332.

Weber, R. P. (1990). *Basic content analysis*. London: Sage Publications.

Westphal, D., & Stern, I. (2006). The other pathway to the boardroom : Interpersonal influence behaviour as a substitute for elite credentials and majority status in obtaining board appointments. *Administrative Science Quarterly*, *51*(2), 169–204.

Westphal, J. D., & Clement, M. B. (2008). Socio-political dynamics in relations between top managers and security analysts: Favour rendering, reciprocity, and analyst stock recommendations. *Academy of Management Journal*, *51*(5), 873–897.

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, *48*(4), 817–838.

Wiersema, M. F., & Zhang, Y. A. N. (2011). CEO dismissal : The role of investment analysts. *Strategic Management Journal*, *32*(11), 1161–1182.

Womack, K. L. (1996). Do brokerage analysts' recommendations have investment value? *Journal of Finance*, *51*(1), 137–167.

Wooldridge, J. M. (2013). *Introductory econometrics : a modern approach* (Fifth edit). Hampshire: South-Western Cengage Learning.

Yu, F. (Frank). (2008). Analyst coverage and earnings management. *Journal of Financial Economics*, *88*(2), 245–271.

Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004). Automating linguistics-based cues for detecting deception in text-based asynchronous computer- mediated communication. *Group Decision and Negotiation*, *13*(1), 81–106.